

Biology versus physics?

There is a tendency for biology to dominate perceptions of science at the expense of support for other disciplines. There is no scientific justification for that situation, which also carries economic dangers.

Paradigm is a problematic word. Thanks to the late Thomas Kuhn, who commandeered it for his own philosophical purposes, it implicitly carries with it the notion that only physicists these days can experience true scientific revolutions. A Kuhnian revolution involves the acceptance of a new conceptual framework for scientific models that is incommensurate with (or cannot be derived by extension of) the traditional framework. The older paradigm turns out to have produced predictions that are approximations of those, closer to observation, that are delivered by its successor.

Whereas physicists can point to Planck and Einstein as classic suppliers of new paradigms, there has always been difficulty in identifying biological conceptual frameworks that fit this definition. Darwin provided a framework where, *pace* various deities, there was little before. The elucidation of the structure of DNA and of the genetic code started a revolution worthy of the title, but those were not incommensurate with any other ideas then accepted in chemistry or biology. The subsequent discovery of non-coding sequences within genes took everybody by surprise but was readily accommodated. Some have proposed Peter Mitchell's chemiosmotic theory (*Nature* 191, 144; 1961) as a Kuhnian revolutionary achievement (*BioEssays* 19, 93; 1996).

Biologists are not alone in occasionally using the word "paradigm" in a narrow context that arch-Kuhnians should thoroughly disapprove of — an improper substitute for "scenario" or, even more specifically, "model". (An example is the hypothesis, described by some as a paradigm, that particular molecules — neurotrophins — are the key to the competition between neurons for survival and to the formation of appropriate inter-neuron connections in the developing brain.)

On the other hand, non-Kuhnian revolutions in biology are often, even usually, sparked by innovations in technique. Some are biological developments — monoclonal antibodies, the polymerase chain reaction — while many other transformational tools have emerged from physics: X-rays, radioactivity, neutrons, positrons, nuclear magnetism, various microscopies, optical tweezers....

Forget about philosophy — that bio-relevance, say some physicists, is a reason to fund physics. How are the mighty fallen, others might respond. From their dominant positions at the heart of the science-industrial-government corpus during the middle decades of this century, physicists are now reduced to justifying their continuing existence on the coat-tails of another discipline.

It is one thing to argue, as this publication has, that disciplines need to look to themselves to produce effective advocacy. It is another to leap onto a superficial notion that any discipline has little more to offer. That dismissive tone comes from the very top, in the recent assertion of President Bill Clinton to the effect that the past 50 years have been the age of physics, whereas the next will be "very likely characterized predominantly as the age of biology". Supporters of the life sciences — some of them scientists who should know better — quote that all too gleefully and glibly.

Physicists and chemists cannot take all the blame for the low public and political appreciation of their disciplines. Who can easily compete with the weekly announcements, trumpeted by the media worldwide, of a new genetic link, a new virus, or another step in the understanding of neural degeneration? And in the eyes of those who look more closely

at the science, there is an appealing grandeur in the unfolding picture of underlying commonality, at the genetic and sub-cellular levels, between the kingdoms of life, let alone species.

Yet take, as one example of contemporary physics in particular, a recent paper that for the first time demonstrated quantum teleportation experimentally (*Nature* 390, 575; 1997). More media-friendly than most "small" physics ("teleportation" was leapt on in many newspaper reports), the result and the underlying principles are nevertheless difficult to comprehend by any standard. Yet for its technical quality, for its intrinsic scientific value within the quantum paradigm that underlies our understanding of light and matter, for its relevance to philosophical paradoxes associated with quantum mechanics, and for its technological potential, such a result, and others that will no doubt follow, gives much reason to celebrate and support the discipline that produced it.

Dangers

Some biologists, benefiting from politicians' justified appeals for increases in health research budgets (see for example page 114), protest that their discipline is not competing with others — as did the chief executive of the US Federation of American Societies for Experimental Biology, saying last week in the *Washington Post* that increases for health research are not intended to be at the expense of other sciences, which are funded from different budget lines. Given the pressures on public funds and the tone set by Clinton, that statement is questionable if not disingenuous.

Despite the dangers of its misrepresentation, Clinton's statement is appropriate as a celebration of biology's increasing relevance to countries' health and wealth. Meanwhile, physicists are rightly attracted by new applications of their work — to probing the structures, properties, dynamics and energetics of complex biomolecules, to non-invasive imaging, and to the analysis and simulation of biological networks and systems.

But, in the United States and in Europe, there is a growing danger that, as the attractiveness and impact of biology develops, an inadequate appreciation of physics will effectively lead to a backlash against it. Perhaps the only allies that physicists have are, for big science, the cultural appeal to the public of fundamental laws of nature and of the Universe. For small physics, as for non-biological chemistry, the principal allies are in industry.

Despite the end of the Cold War, defence and civil industrial interests in "small" science are as strong as ever. Yet attempts to marry industrial and scientific perspectives — such as the technology foresight exercises now carried out in several countries — fail to generate political resonance. Nevertheless, new chemicals, advanced materials, telecommunications, semiconductors and other electronic, magnetic and photonic technologies have major industrial relevance, as much significance for national economies as ever and an increasing impact on society. The fact that the results of physics in particular have a way of unobtrusively and unpredictably invigorating apparently unrelated disciplines and technologies, even to a revolutionary extent, is an additional benefit. As skill shortages grow, should not all that be something for politicians to worry more about? □



NASA lines up for a bigger slice of the biological research pie

[WASHINGTON] The US space agency NASA is bracing itself for the receipt by the end of this month of an avalanche of proposals from research groups wishing to join its new Astrobiology Institute for the interdisciplinary study of life in the Universe.

If successful, the experimental institute could take the agency's research activities beyond its traditional roots in fields such as astrophysics and planetary geology. This 'virtual institute' — which NASA planners envisage as a semi-independent network of organizations working on similar research problems around the country — will be 'hosted' by the Ames Research Center outside San Francisco, where a director and a few staff will provide computing and administrative support and overall programme direction.

The institute will pioneer techniques in scientific collaboration, using the next-generation Internet to let researchers collaborate via video links, and share expensive instruments and computer simulations. Funding for the first year will be about \$7 million, climbing to \$8 million next year and \$10 million thereafter. NASA says it is committed to the programme for "at least 20 years".

The scope of research at the Astrobiology Institute will be broad, ranging from investigation of the origin of life on Earth, through the question of how terrestrial ecosystems evolve to the search for life on other planets. The planners hope that the interdisciplinary study of these problems will, in effect, create a

branch of science — astrobiology — whose exact definition is not yet established. "We don't really know what this field is going to be," says Gerald Soffen of the NASA Goddard Space Flight Center in Maryland, who has been asked by the agency's administrator Dan Goldin to help shape plans for astrobiology. Soffen was project scientist for the Viking Mars missions of the 1970s, and later served as the agency's director of life sciences.

Interest in the subject was stirred last year by the report of signs of fossil life in a martian meteorite. NASA and the National Science Foundation have already invested extra funds for research on life in extreme environments on Earth and NASA is to introduce a research programme in evolutionary biology this year, with \$4 million going to individual scientists working on questions of life's origins.

The Astrobiology Institute has strong support from Goldin, who has been pushing NASA to beef up its biology programme for more than a year (see *Nature* 384, 601; 1996). Goldin has made a point recently of meeting biologists, including David Baltimore, the new president of the California Institute of Technology.

The relative wealth and political support enjoyed by the biological sciences is one reason for NASA's interest in the field. But Goldin also wants his agency to be a leader in technological innovations such as computers based on DNA's ability to organize information, and the creation of autonomous, 'think-



Hard evidence: this martian rock has already had a major influence on thinking at NASA.

ing spacecraft' based on biological principles.

Today, Goldin says, NASA doesn't have the "intellectual underpinnings" for such work. "The biological revolution has passed the space program by," he lamented last month at a meeting of NASA's external advisory council. Goldin's answer is to raise both the profile and the number of biologists in NASA. He hopes to hire a biologist as the agency's next chief scientist, a hitherto mostly symbolic post that has been vacant for months.

Goldin also told the advisory council he would like to see spending on the Astrobiology Institute eventually reach \$100 million a year. "You just wait for the screaming from the physical scientists [when that happens]," he said.

But the institute will start modestly. NASA hopes to pick between seven and ten institute partners in April, who will receive funding for three to five years. Michael Meyer, astrobiology discipline scientist at NASA headquarters, says the primary reason for choosing institutional members rather than simply funding individual researchers is to take advantage of the interdisciplinary links that member organizations such as universities will bring with them. Institute members are also expected to contribute resources beyond the funding that NASA provides.

Although the NASA advisory council supports the institute as a worthwhile experiment, some members at last month's meeting expressed unease about the concept's lack of definition. Many believe success will depend heavily on the leadership of the institute's director, who has yet to be chosen. Soffen says the agency is searching for a "non-NASA person who has a vision".

Interest already seems high in the research community. NASA received more than 70 letters of intent to respond to the solicitation for institute members. That prompted Soffen to worry that the agency will not have enough money to satisfy the demand. "It kind of worries me what we turned on," he said.

Tony Reichhardt

India and United States to share satellite data

[NEW DELHI] After more than a decade of refusing to do so, India has agreed to share with the United States weather and atmospheric data gathered by its INSAT geostationary satellites over the Indian Ocean on a "near real-time basis". But it has stipulated that the United States should not disseminate the Indian data to a third country without its consent.

India has three INSAT satellites collecting data, with a resolution of 2.75 km in the visible band and 11 km in the infrared. INSAT-2E, to be launched this year, will have improved resolution, as well as the ability to measure

atmospheric moisture.

The agreement, signed last week in Washington, represents a victory for US scientists, who have been pressing since 1987 for access to INSAT data to help them fill a data gap over the Indian Ocean.

The latest agreement — reached after intense negotiations between NASA and the Indian Department of Space (see *Nature* 375, 441; 1995) — marks a policy shift that ends a long-standing irritant in Indo-American scientific collaboration.

Under the agreement, initially planned for five years, NASA and the US National Oceanic and Atmospheric

Administration (NOAA) can access high-resolution processed INSAT data via a dedicated link between NASA in Washington DC and IMD in New Delhi.

Processing involves only a few seconds, so US agencies will receive INSAT data in "near real time", says an IMD official. India is to get direct access to NASA's Earth and atmosphere science databases, including data from NOAA's Geostationary Operational Environmental Satellites (GOES) over the Pacific Ocean. IMD says this would help to sharpen its forecasts of monsoons and tropical cyclones.

K. S. Jayaraman

French researchers oppose biomedical reform plan

[PARIS] Claude Allègre, the French minister for national education, research and technology, has put forward proposals for the reform of the national biomedical research agency, Inserm, aimed at streamlining its administration and increasing its capacity for strategic planning.

But the moves are being viewed with concern by some Inserm researchers, who feel that they would lead to a reduction of their influence over research strategy. More than 200 laboratory heads and other researchers have signed a petition protesting against the proposals for reform of the agency.

One of the main reforms being proposed by Allègre is that there should be a strengthening of the role of the agency's administrative council in making strategic decisions on research directions that have in the past been agreed jointly by the scientific community and Inserm management.

Some scientists have interpreted this as an attempt by the ministry to seek greater control over the agency, and are concerned that the relative lack of scientific experience among members of the new board might lead to a lack of adequate focus in research goals.

Evidence of an apparent desire for greater centralization also comes from plans by the ministry to take direct control of more than half of the funds for laboratory equipment and supplies, which were previously overseen by scientific committees.

At the same time, the ministry proposes to reduce the number of scientists appointed to the agency's elected bodies for evaluating research, while the scientific board will lose its powers for setting research strategies and choosing recruits. Both are being seen as weakening the influence of researchers over broad research strategy.

The fourth major reform outlined is the division of Inserm into scientific departments — similar to the Centre National de la Recherche Scientifique (CNRS) — each of which will be headed by an official appointed by the ministry. The move has surprised scientists as Inserm, with 2,115 scientists and 2,755 technicians, is relatively small compared with CNRS, which has more than 26,000 staff; indeed the whole of Inserm is smaller than the CNRS's life sciences department.

The petition expresses the concern that the moves seem aimed at converting Inserm from an independent research agency with control over its own research activities to a research council distributing funds to university groups and hospitals on the basis of policies defined by the ministry.

The scientists who signed the petition claim that the ministry has failed to consult either Inserm's scientific boards or its trade unions over the planned reforms, describing this as a "dirigiste, centralist, and technocratic" attitude.

This accusation is refuted by the ministry, which claims that Inserm representatives were fully consulted. But Nicolette Farman, from the National Union of Scientific Researchers, says that, while the trade unions were asked for their opinions of the proposed changes, they were excluded from the preparations of the reforms themselves.

In a thinly-veiled attack on the biomedical research agency, the ministry said in a statement that, while its scientific development over the past few years had been globally "positive", it lagged behind in fields such as "instruments, the development of drugs, physiology, clinical trials and medical information", justifying the proposals for reform.

Eric Glover

Orbital space debris 'poses main threat to shuttle crew'

[WASHINGTON] The US government urgently needs to develop an interagency surveillance strategy involving both the Department of Defense (DoD) and the National Aeronautics and Space Administration (NASA) to combat the growing problem of space debris, according to the General Accounting Office (GAO), the investigative arm of Congress.

The recommendation, made in a report published by the GAO last month, coincides with a separate report from the National Research Council (NRC) urging the space agency to boost its efforts to protect the space shuttle from meteoroids and orbital space debris. The council says these now represent "the single greatest threat to the shuttle and its crew" — even higher than the hazards encountered during launch and ascent.

The United States already has a sophisticated Space Surveillance Network, operated by the US Space Command, on which both the DoD and NASA rely for information on objects orbiting the Earth. But the GAO says that this is not capable of providing the information the space agency needs — in terms of both locating and identifying the size of such objects — adequately to predict possible collisions with multibillion-dollar space programmes, such as the international space station.

Although the defence department has plans to modernize its surveillance network radar system and to develop three new ballistic missile warning systems, the GAO says that these plans "do not adequately consider DoD's or NASA's surveillance requirements". Nor does the National Space Policy, published in 1996, make provision for an interagency mechanism to ensure that NASA's requirements are met.

The NRC report was drawn up by a panel set up by its Aeronautics and Space Engineering Board committee to address the potential threat to space shuttle missions posed by orbital debris such as spent rocket bodies, satellite fragments and even paint chips. It points out that NASA's four operating shuttles, which were designed in the 1970s, were not built to repel bombardment by such debris, as it was not recognised at the time to be a substantial threat.

But the growing amount of debris over the past decades, says the committee, has almost doubled the likelihood that the crew might be harmed or that the spacecraft might incur major damage. The committee also points out that more than 95 per cent of the debris that could critically damage the shuttle is too small to be picked up by current ground-based sensors operated by the DoD. □

Gould to be next year's president of AAAS

[WASHINGTON] One of science's best known names, Stephen Jay Gould, has been chosen as the next president-elect of the American Association for the Advancement of Science (AAAS).

Gould, a Harvard University palaeontologist and evolutionary biologist as well as a best-selling author, assumes the post on 18 February, at which time M. R. C. Greenwood, a former associate director of the White House science office and now chancellor of the University of California at Santa Cruz, will become president. Gould will then succeed her in January 1999.

Gould is best known scientifically for his work in developing the theory of punctuated

equilibrium, which holds that evolutionary change occurs in sporadic bursts rather than at a steady, gradual pace. He has also written 15 books on science for general readers, and is expected to emphasize public understanding of science during his tenure.

"I want to make people less scared of science so they won't see it as arcane, monolithic and distant, but as something that is important to their lives," Gould said in a statement following the announcement of his election last month.

"We have to ask ourselves why so many people have become so afraid of science — it is no more difficult to learn than other things."

Tony Reichhardt

Japan maintains research funding boost...

[TOKYO] Despite an austere budget that reflects Japan's struggling economy, the government has yet again agreed to increase science and technology spending for the 1998 fiscal year, beginning on 1 April. Overall science expenditure across all ministries will increase by 4.9 per cent, with particularly increased support for basic science.

The overall increase in the research budget is relatively modest compared with past increases. But science has fared reasonably well considering that public spending for other areas, such as agriculture, forestry and fisheries and overseas development assistance, has plunged dramatically.

The increase in science spending is also in line with Japan's five-year plan for science and technology, under which the government has promised to increase spending on research by 50 per cent between 1996 and 2001.

The total government budget of ¥77,670 billion (US\$595 billion), an increase of 0.4 per cent from the budget for the current fiscal year, is the smallest increase in recent years. The figures, which were approved by the cabinet at the end of December, have yet to be approved by the Diet (Japan's parliament) in its regular session, which starts on 12 January.

This is the first budget compiled under the terms of the Fiscal Structural Reform Act, which came into effect in November. The act seeks to cap major spending commitments in

order to hold government deficits to 3 per cent of gross domestic product by 2003, and encourages budgeting restraint.

But science spending, particularly in basic research, has increased. Life sciences have done particularly well, with brain- and genome-related research gaining substantial budget increases (see table). This is in line with last year's opening of the Brain Science Institute and this year's planned opening of the Genome Frontier Research Institute, both backed by the Science and Technology Agency (STA) and the Institute of Physical and Chemical Research (RIKEN).

STA's total budget, however, has fallen by 0.2 per cent largely because of cuts in support for nuclear research. This is the first time in STA's history that nuclear research has not ranked top in terms of spending. Basic research, with an 11.8 per cent spending increase, becomes the largest category in STA's budget. Spending on nuclear research will fall by 4.9 per cent, and the budget for the Nuclear Power and Fuel Corporation, hit by scandals over nuclear accidents, will be cut by 10 per cent.

The overall budget for the Ministry of Education, Science, Sports and Culture (Monbusho) has decreased by 0.4 per cent, as a result of cuts in support for education, but its expenditure on scientific research has increased by 5.1 per cent. Postdoctoral fellow-

Highlights of Japan's science and technology budget 1998 (in billion yen: US\$1 = 132 yen)

Science and Technology Agency (STA)		
Total budget	733.1	-0.2%
General R&D budget	581.1	+1.7%
Space	17.6	-1.7%
SPRING-8	15.6	-19.1%
Human genome	6.7	+162.5%
Brain research	13.3	+33.7%
STA fellowships	3.6	+16.1%
Nuclear power	76	-11.7%

Ministry of Trade and Industry (MITI)

Total R&D budget	492.7	+4.3%
Agency for Industrial Science & Technology	169.2	+4.4%
New Industry R&D	30.0	+6.8%
Joint research with universities	2.2	New
Human Frontier Science Programme*	4.2	+5%

*includes 2.6 billion yen from STA

Ministry of Education, Science, Sports and Culture (Monbusho)

Grants in aid for scientific research	117.9	+5.1%
Japan Society for Promotion of Science	41.5	+8.1%
(postdoctoral fellowships)	(16.1)	(+11.6%)
Joint research with industry	106.5	+5.0%
Academic information network	36.4	+0.4%
Centre of excellence programme	12.7	+12.8%

ship schemes at Monbusho, aimed at increasing the number of Japan's postdocs to 10,000 by the end of the decade, have been given increased funding. But national universities face budget cuts, resulting in a substantial increase in university tuition fees from 1999.

Both Monbusho and the Ministry of Trade and Industry (MITI) will receive increased spending to promote collaborative research between universities and industry.

One area to receive a marked increase in spending is the drive to curb global warming, following the adoption of the Kyoto protocol which sets targets for reducing greenhouse gases (see *Nature* 390, 649; 1997). At the Kyoto conference last month, Japan agreed to reduce greenhouse-gas emissions by 6 per cent from 1990 levels between the years 2008 and 2012. The Environment Agency will receive increased spending for efforts to reduce its greenhouse gases, and MITI will be allocated funds for reducing energy expenditure and for more research into alternative energy sources.

The budget plan's critics claim that reducing greenhouse gases to the required level will mean that Japan's energy production must become more dependent on nuclear power. This, says one MITI official, will require 20 additional nuclear-power reactors to be built. But with considerable cuts in spending on nuclear power, many believe that would be impossible.

Asako Saegusa

... as science escapes cuts in South Korea

[TOKYO] Recent stock market and currency turmoil has hit South Korea's economy hard, but despite this the government has approved a five-year science and technology innovation plan designed to increase government-funded research and development to boost economic growth.

Under the plan, approved last month by a senior government committee headed by the deputy prime minister, South Korea will increase the government expenditure allocated to R&D from 3.9 per cent to at least 5 per cent by 2002.

The government had previously planned to increase its total expenditure by 5.7 per cent next year, with R&D expenditure increasing by 12.3 per cent to 3,088 billion won (US\$1.82 billion). But the International

Monetary Fund has asked for a 10 per cent cut in total government expenditure as part of its financial rescue package, and at least 10 trillion won is expected to be cut from the 1998 budget.

Government officials and academics expect infrastructure and construction budgets to bear the brunt of the cuts, according to Yong Hwan Kim, director of the policy planning division of the Ministry of Science and Technology.

The new five-year plan follows on the heels of the special law on science and technology innovation approved last July. The law was a response to severe criticism of South Korea's R&D policy, which was said to be grossly inadequate for fostering innovation.

The plan is designed to stimulate innovation-led

economic growth and industrial development by modernizing South Korea's science and technology capability, bringing it up to the level of advanced industrialized countries.

Kim acknowledges that this will be "quite difficult" in the current economic climate, but argues that improving and increasing basic research is essential for future prosperity.

The academic community has welcomed the plan. But researchers will not escape the economic upset completely unscathed. The dramatic fall in South Korea's currency is making it increasingly expensive to buy scientific equipment from overseas.

The general budget cuts will be decided by the National Assembly next month.

Richard Nathan

Universities seek to atone for Nazi past

[MUNICH] A university from the east has for the first time joined moves by some of Germany's older universities to address a murky aspect of their Nazi pasts. Rostock University on the remote Baltic coast is to search its archives for people whose academic qualifications were removed on racial or political grounds.

The university is planning to rehabilitate those who suffered in this way, even if in most cases this will have to be done posthumously. The first steps were discussed at the university last month.

The problem originated in 1933, when the Nazi government issued a law that stripped those who had decided to leave Germany because of persecution not only of German citizenship but also of academic qualifications, mostly doctorates.

Soon after, the law was extended to any German resident exhibiting 'antisocial behaviour' — a move targeted at Jewish, communist and dissident academics.

Each university was ordered by the min-

istry of education to alter its rules to facilitate the derecognition of doctorates. Although the 31 universities then in Germany differed in the extent to which they applied the law, by 1945 an estimated 1,000 academics had lost their titles in this way.

After the war, partly in a bid to forestall intervention by the occupying Allied forces, universities in the new West Germany were quick to annul all decisions based on fascist ideology made in the Nazi era, including the derecognition of academic qualifications.

The 24 universities in the new West Germany each decided to restore academic qualifications to those from whom they had been stripped. But this was done only at the request of individuals whose cases would then be examined.

In contrast, the socialist regime in East Germany chose to build completely afresh, rather than reversing fascist laws and decisions. East German universities therefore did not feel it necessary to address the issue of derecognized qualifications.

Even in the West, however, movement was initially slow. In 1947, half of all West German professors were former members of the Nazi party. Indeed, a former chairman of a Nazi military court was even appointed rector at Marburg University. As a result universities were reluctant to look critically at their activities during the Third Reich.

Recently, however, this has changed. "The postwar generation started taking over the top positions in universities in the 1980s and 1990s," says Eckart Krause, a researcher in the faculty of history at Hamburg University. "This meant that universities gradually adopted more open-minded attitudes to their pasts."

In 1991 Hamburg became the first university in Germany to publish a report on its role in Nazi times. This included an official apology for the stripping of academic qualifications, and a statement declaring that the university now "derecognizes the derecognition" of qualifications.

Other universities, including Munich,

Space 'time capsule' could send a message to the future

[PARIS] A French artist is hoping to launch a satellite into orbit early next century that would return to Earth in 50,000 years, delivering a cargo of works of art and compact discs containing a contemporary 'library of Alexandria', as well as millions of messages from today's inhabitants of Earth. Its heat shield would create an artificial aurora borealis on re-entry, to alert the then inhabitants of the Earth to its arrival.

The craft, to be called Kéo as it comprises the only three sounds common to all languages, is the brainchild of the Paris-based artist Jean-Marc Philippe. Kéo, the 'archaeological bird', has been developed in collaboration with research and space organizations including the French Atomic Energy Commission, the Aérospatiale company, and Sup'Aéro, the Toulouse-based *grande école* for aeronautics and space.

The same organizations, as well as the European Space Agency, are also backing another project of his to send the first sculpture to another planet, Mars. Both collaborations involve no direct funding, being developed through 'in kind' contributions of materials and time.

The French space agency CNES is seeking to arrange a free launch of Kéo on Ariane-5, according to Josette Runavot, deputy head of the agency's planetary exploration department, who points out that Ariane-5 is able to offer 'piggy-back' flights for small satellites alongside commercial payloads. The project is "marvellous", she says.

After establishing himself as a painter in



Round trip: Kéo, described as an 'archaeological bird', would return to Earth in 50,000 years time.

the 1970s, Philippe — who has a PhD in astrophysics — turned to developing new technologies to artistic ends. Through joint ventures with the US company Raychem and French research agencies, he has pioneered the development of 'shape memory alloys', compounds invented for use in US fighter aircraft that are able to 'memorize' precise shapes at various temperatures.

By adapting new technologies, artists can broaden creative possibilities, says Philippe. His work with alloys has yielded a copper-zinc-aluminium sculpture, 'Hermaphrodite', a bust of a male Greek torso at 20 °C that changes to a woman's torso at 55 °C.

Kéo will have such sensitive alloys in its solar panels so they move up when the craft is in sunlight and down when it is in Earth's shadow, to give the appearance of a large

gliding migratory bird. Inside the satellite samples of sea water, soil, air and a drop of human blood are held in a diamond, the surface of which is engraved with a highly conserved region of human DNA.

The capacity of the craft's titanium nitrate-plated glass discs is sufficient to store four pages of "uncensored" text from every inhabitant of Earth, says Philippe, who says the main aim of the project is to invite as much of humanity as possible to reflect on its history and destiny. He is preparing an international effort to collect text, and is in discussions with sociologists as to how the compilation could be exploited for research purposes.

For the second project, European space scientists have offered to trim the size of their instruments to make room for a threadlike sphere 10 cm in diameter containing a torus of shape-memory alloys that cause the sculpture to close during the cold martian night and open during the day to reveal a pyramidal diamond at its core, engraved with the DNA double helix. Philippe has submitted the sculpture to NASA's call for proposals for its Mars flights. But Marcello Coradini, coordinator of Solar System missions at the European Space Agency, says that if the bid is unsuccessful, his agency "will do its utmost" to fly the sculpture on its Mars Express probe.

Coradini argues that having a camera sending back images of a reactive artwork from Mars could be of "huge public interest".

Declan Butler

Freiburg, Kiel and Frankfurt, have now also started going through their files to check for derecognition of doctorates and injustice in general in Nazi times.

Each university is proceeding in a different way, but all are checking every case individually. "We cannot and we do not grant a general rehabilitation," says Christian Winter, the vice-president of Frankfurt University.

"It is most important to discriminate between those who lost their titles on racial or political grounds and those who lost their titles on genuine criminal grounds."

Universities have found that it is not always easy to discriminate between the possibilities. Some cases are clear. For example, if the grounds for divestment of title were given simply as homosexuality or disloyalty to the Nazi regime, no longer crimes today, then rehabilitation would be automatic.

But if an academic was stripped of his or her title because of a financial offence, then the exact nature of the offence becomes important. Genuine criminal grounds were not unusual. Of 59 persons stripped of doctorates by Hamburg University, for example, eight were involved in genuine crime, including murder.

In the eastern part of Germany, some universities remain unaware of the issue — Rostock learned of its involvement only after being contacted by *Nature*. Other universities say they lack the resources to carry out time-consuming investigations.

For most of those affected, the efforts come far too late. Many have died, and it is difficult to track down those who are still alive. Aware of this, Munich University, in its 1996 statement revoking its own derecognition of 135 academic titles, described it as "first of all a gesture for the bereaved" with which the university "faces its responsibilities arising from its history". **Matthias Strobl**



Honour regained: the author Thomas Mann was stripped of his honorary doctorate by the University of Bonn in 1936. It was returned to him on the university's initiative in 1946.

Nobel laureates in bid to revamp science teaching

[SAN FRANCISCO] A group that includes ten Nobel laureates will learn shortly whether it has been successful with its offer to restructure academic standards for California's science curriculum. The move follows a decision last month by a state commission to reopen a process under which the group's bid had earlier been rejected.

The Nobel prizewinners include such well-known names as David Baltimore, Glenn T. Seaborg, Henry Taube, Dudley Herschbach and Paul Berg, and the group is keen to counteract what its members regard as a 'dumbing down' trend in science education. It seeks to emphasize fundamental concepts and to ensure that schools produce competent students who will go on to science studies in colleges and universities.

The group says it would insist on well-written textbooks, scientifically trained teachers, better salaries for instructors, and a focus on basic scientific concepts and mechanisms. It suggests that students should begin physics in ninth grade (at 14 years of age), for example, and graduate with an understanding of Newton's laws of dynamics, elasticity, linear systems, electromagnetic mass, the laws of thermodynamics, and quantum behaviour, among other topics.

The group also proposes that students should start chemistry lessons in fifth grade and learn the periodic table and classification of elements, solubility, chemical equilibrium, atomic structure, chemical kinetics and thermochemistry, as well as other topics. In biology, it says, the emphasis should be on genetics and molecular biology.

The group argues that vague approaches to teaching science that fail to challenge students turn them off science and help create an environment where, for example, only 3 per cent of men and 1 per cent of women graduating from US universities have degrees in computer science.

Over the past decade, the number of students graduating in computer science in the United States has dropped by nearly a half, despite burgeoning job opportunities. Scientists must step forward to improve science education and to help lure more financial backing for better teaching laboratories, according to the group.

"(Scientists) are equipped and motivated to insist on teaching fundamentals and basics," says Seaborg, professor of chemistry at the University of California at Berkeley and a former chairman of the Atomic Energy Commission. Seaborg says that science deserves special attention in schools because so much of the job market now requires scientific or technological expertise.



Seaborg: 'go back to basics in teaching.'

"Technology firms are hurting because they can't find people with elementary scientific knowledge," he says; even those in other fields need a basic understanding of science in order to perform adequately as citizens. Seaborg was the lead author of a

1983 report, *A Nation at Risk*, which detailed educational weaknesses and stimulated a wave of reforms.

The group, which calls itself Associated Scientists, was put together by Shoumen Datta, a molecular biologist who has been active in various education initiatives. Datta says he hopes that curriculum reform in California will stimulate a transformation of teaching methods throughout the country, helping to instil a general standard of excellence in US schools. "We are against the 'process, process' mentality in schools," he says, instead of a focus on accurate content.

Associated Scientists resubmitted its offer to shape academic standards, which it has said it is prepared to do at no cost, shortly before Christmas after a protest had forced the state commission for academic content and standards to rescind the offer of a contract made to another group.

The Institute for Science Education at California State University, San Bernardino, made up primarily of professional educators, had originally won the assignment with a bid costing \$178,000. One member of the commission's assessment team had explained its decision to award the contract to this group by saying the Nobel laureates "wouldn't know a classroom if you put it in front of them".

Marci McFadden, spokeswoman for the commission, subsequently argued that Associated Scientists had not included sufficient details about the education experience of its members, or the mechanisms by which they planned to accomplish their work. But when the State General Counsel informed the commission it had acted improperly in offering the contract to the San Bernardino institute, the bid process was reopened.

Final proposals were received in the week before Christmas, and the award to the winning contractor is due to be announced by 15 January. It is widely expected that the two teams will collaborate on the project, combining the San Bernardino group's expertise in education techniques with the Nobel laureates' knowledge of science. **Sally Lehrman**

Sequencing funds 'must remain focused'

[WASHINGTON] Some of the leading US researchers involved in genome sequencing under the Human Genome Project have urged their major source of funding not to become too distracted with other activities which, they argue, could slow progress towards completion of the project.

Principal investigators funded by the National Human Genome Research Institute (NHGRI) to carry out large-scale sequencing made their comments shortly before Christmas at a workshop with members of an NHGRI advisory council subcommittee that is planning strategy for completing the sequencing of the human genome.

The meeting also addressed other thorny issues involved in efficiently completing the sequencing by the target date of 2005. The concerns of both institute planners and researchers include recruiting and training dozens of investigators every year, and keeping those they have in the face of attractive offers from industry.

Institute officials, led by Francis Collins, the director of NHGRI, laid out a scheme that projected institute spending of \$60 million a year to finance the sequencing of 60 per cent of the human genome by 2005. (The other major funding body is the Department of Energy.)

But some investigators suggested that the institute is not being sufficiently ambitious. Robert Waterston, director of the Genome Sequencing Center at Washington University in St Louis, Missouri, referred to NHGRI's planned investment of \$60 million out of a total extramural grant budget of \$145 million in 1998, and said that the figure indicates "a significant amount of distraction" by the institute.

Waterston said that "it's possible to spend too much money too fast and cause people to become inefficient," especially as the project expands. But the institute, as the sole funder of genomic sequencing at the National Institutes of Health (NIH), had to be "very careful to maintain focus". For instance, he said, it should ensure that the funding for applications of genome sequencing is shared by other NIH institutes.

In response, institute officials said that the spending plan represents only a 'baseline' for planning purposes. Collins says that, although he is unable to promise money from government budgets not yet written, "the sequencing part of the genome project is our highest priority. Nothing else assumes pre-eminence."

Collins says the money spent by the institute on human sequencing has more than

In view of your Biology degree, I'm commuting your jail sentence to 3 years of Human Genome sequencing.



tripled in two years, from less than \$20 million in 1996. At the same time, he says that other "very important" projects deserve the support they are receiving. These include projects to advance sequencing technology, to develop databases that will allow the efficient use of sequence information, and to create a public catalogue of variations in human DNA known as single nucleotide polymorphisms.

"If we need to diminish [non-sequencing activities] to make sure we get the sequence done, we will," Collins says. But he expressed

uncertainty about how the question of funding would develop over the next two or three years.

Collins says that gauging future costs is difficult as it is not clear how quickly the US sequencing laboratories will gain capacity. Nor can it be predicted how quickly or how far sequencing costs will fall. But "no one should make the mistake of assuming that we have locked in this \$60 million a year and that's all there's going to be. If we have to find more, we will."

With fewer than 3 per cent of the genome's 3 billion bases sequenced so far, institute officials are also facing the complex challenges of recruiting scientists and of increasing at least tenfold the country's sequencing capacity in order to complete the job in time.

Eric Lander, director of the Whitehead Institute at the Massachusetts Institute of Technology, told last month's meeting: "I'm concerned that, the way the major sequencing is going, we're not going to be attracting any new young blood."

Those present at the meeting debated strategies for establishing sequencing centres and expanding existing ones while improving quality and efficiency. Another problem discussed was how to cost accurately the sequencing of a base pair to allow budget planning and comparisons of competing grant applications.

Meredith Wadman

Indian guidelines allow limited gene screening

[NEW DELHI] The Indian Council of Medical Research has released draft ethical guidelines on biomedical research. These would allow genetic screening in employment with the consent of employees, but prohibit life insurance companies from making tests a prerequisite for insurance.

The guidelines would also allow medically terminated fetuses to be used for transplanting to "patients for whom no other form of treatment is available". But a strict ban on animal-to-human transplants is proposed.

The panel, headed by M. N. Venkatachalaiah, a former chief justice of the supreme court, has updated guidelines issued in 1980. It has included regulations for research into areas such as human genetics, organ

transplantation and assisted reproduction.

Employee screening is justified for genetic disorders that might jeopardize the safety of others. Thus airline companies can screen pilots for sickle-cell anaemia, which can affect an individual's actions when "exposed to atypical atmospheric conditions", endangering the lives of passengers.

But family members should not be entitled to know one another's genetic diagnosis because in India "revealing the information that the wife is a carrier of a recessive disease may lead to the husband asking for a divorce". But if a person has AIDS, it should not be kept secret, the report says.

Genetic screening is at present voluntary in India (although screening newborn babies at risk may be made compulsory). Prenatal

diagnosis is permitted only where it is relevant to the health of the fetus or mother – not to select for sex.

The guidelines allow research on human embryos up to 14 days old and also the creation of 'abnormal' embryos using sperm and eggs taken from high-risk parents to study the transmission of genetic disorders.

But the transfer of any manipulated embryo to a human uterus, and the commercial exploitation of embryo research, would be banned.

All international collaborations involving human genetic research would be subject to regulations laid down by the Indian government. No DNA samples are allowed to leave India without adherence to previously established guidelines.

K. S. Jayaraman

Clinton 'will seek extra \$1bn for biomedical research'

[WASHINGTON] President Bill Clinton is reported to have decided to seek a substantial increase in the budget of the US National Institutes of Health (NIH) of \$1 billion, or nearly 8 per cent, in the budget request for 1999 that he plans to send to the Congress next month.

White House officials are said to have told *The New York Times* that Clinton has given his assent to the increase in the budget of the \$13.65-billion-a-year agency. Only last month Clinton declared that the next 50 years "will very likely be characterized predominantly as an age of biology".

Leading members of Congress have already indicated their intention to seek a substantially larger increase for the NIH. Congressman John Porter (Republican, Illinois), the chairman of a key House of Representatives spending subcommittee, has repeatedly called for a boost of at least \$2.5 billion for NIH in 1999 (see *Nature* **390**, 6; 1997). And in November the Speaker of the House, Newt Gingrich, urged a doubling of spending on biological research (see *Nature* **390**, 321; 1997).

The 8 per cent increase would mark a departure for Clinton, who in 1997 and 1998 asked for NIH increases of 3.9 per cent and 2.6 per cent respectively; Congress granted increases of 6.9 per cent and 7.1 per cent.

Germany goes Dutch for climate change research

[MUNICH] The Netherlands and Germany are to collaborate in research on climate change and sea and ocean management. Scientists from the Netherlands Institute for Sea Research, on the North Sea island of Texel, and from four research institutes in Bremen, on Germany's North Sea coast, are to co-operate in a project called NEBROC (Netherlands Bremen Oceanography).

The joint venture was approved by the science ministers of the Netherlands, Germany and Bremen at a meeting last autumn on the German polar research vessel *Polarstern* (*Polar Star*).

Funding for the research programme will come equally from both countries, with Germany and the Netherlands having each agreed to donate Dfl800,000 (US\$443,700) a year for the next five years. The institutes will share their facilities, including research vessels, and will set up a college — the European Graduate College for Marine Sciences — at which 21 young scientists will research areas such as climate history and oceanic carbon cycles.

UK funding boost for laboratory equipment

[LONDON] Britain's Wellcome Trust has linked with the Higher Education Funding Council for England (HEFCE) to launch an initiative to provide funds for research equipment in universities. A new Infrastructure Panel has been set up by the trust to channel £15 million (US\$25 million) a year into support for large equipment and refurbishment grants; applicants for the latter must already be in receipt of "substantial trust funding". The panel will also administer a Joint Equipment Initiative with the HEFCE.

Under the joint initiative, the funding council will provide half of the funding agreed by the trust for some items of equipment costing between £200,000 and £2 million. According to Brian Fender, the chief executive of HEFCE, the partnership will help to develop "a more strategic approach to infrastructure funding".

German company set up to exploit genome data

[MUNICH] A new biotechnology company, Artemis, is to be set up in Germany on the initiative of Christiane Nüsslein-Volhard, director of the Max Planck Institute for Evolutionary Biology in Tübingen and winner of the 1995 Nobel prize for medicine. Artemis's research will concentrate on possible medical applications of Nüsslein-Volhard's findings about the zebrafish genome and similar results by Klaus Rajewski, head of the department of immunology at the University of Cologne, from his work with mice.

Both Nüsslein-Volhard and Rajewski will keep their present research positions, and will act only as advisers for Artemis. Headed by Peter Stadler, a former research head of Bayer, Artemis will cooperate with the US company Exelixis, based in San Francisco. In the spring, six scientists — four of whom are former members of Nüsslein-Volhard's and Rajewski's research groups — begin research in laboratories in Tübingen and Cologne.

\$1.8m pay-out over radioactive oatmeal

[WASHINGTON] The Quaker Oats Company and the Massachusetts Institute of Technology (MIT) have agreed to pay \$1.85 million to settle a lawsuit involving the feeding of radioactive oatmeal to boys in the 1940s and 1950s. The class action lawsuit was brought by several dozen plaintiffs among more than 100 who participated in the experiment.

The boys had been fed cereal containing radioactive iron and calcium in an experiment seeking to trace the absorption

of the oats throughout the body. MIT, whose president apologized for the experiment in 1994, issued a statement last week pointing out that no discernible health effects were found by a state task force that reported in 1994. But lawyers for the plaintiffs said that the consent form used in the experiment neglected to mention that the oatmeal contained radioactivity.

Call for a shake-up at Robert Koch Institute

[MUNICH] In one of its most critical assessments, Germany's science council, the Wissenschaftsrat, has recommended that the historic Robert Koch Institute in Berlin be restructured from top to bottom. The institute was founded in 1891 as Germany's centre for control of infectious diseases. But it has recently been growing in an apparently haphazard way, forming loose attachments to other institutes rather than absorbing them. The newest attachments are the German AIDS Centre and the Institute for Epidemiology and Social Medicine.

The Wissenschaftsrat has now recommended that the management structure of the institute be streamlined and that the institute be restricted to its original brief of handling infectious disease. It also suggests that the variable quality of research carried out at the institute be improved by focusing on fewer themes and by forcing research groups to compete more for institutional money.

Aharonov and Berry win Wolf physics prize

[LONDON] An Israeli and a British physicist have shared this year's Wolf Prize in physics, awarded by the Israeli-based Wolf Foundation, for "the discovery of quantum topological and geometrical phases" — changes to the phase of quantum-



mechanical wavefunctions as a result of topological and geometrical transformations — "and their incorporation into many fields of physics".



Yakir Aharonov (top left), who holds teaching posts at the University of South Carolina in the United States and Tel Aviv University, and Sir Michael Berry (below left), professor of physics at the University of Bristol, will share the \$100,000 prize.

Both have their names associated with important phenomena in quantum physics: Aharonov was the co-discoverer with the late David Bohm of the 'Aharonov-Bohm effect', and Berry is known for the 'Berry phase'.

Hawksbill turtles still endangered

Sir— In a recent Commentary article¹, Nicholas Mrosovsky criticized the International Union for the Conservation of Nature (IUCN) for the lack of documentation supporting the listing of the hawksbill turtle (*Eretmochelys imbricata*) as 'critically endangered' in the 1996 IUCN *Red List of Threatened Animals*².

The listing of the hawksbill as critically endangered was based on a rigorous evaluation by IUCN Marine Turtle Specialist Group members involved in preparing the status justification. I am a member of that group, and a hawksbill specialist. Population declines of 80% over three generations are evident or can be inferred in all ocean basins throughout the species' circumtropical range. The species has been listed by IUCN as endangered since 1968.

Publication of a summary of all the data considered in the listing process is time-consuming, and our working group (all volunteers) is moving forward as fast as possible. It should be pointed out that all the documents we are reviewing are freely available in the literature, including a 601-page review of the hawksbill's status³ sponsored by the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES).

An unfortunate effect of Mrosovsky's Commentary will be to impart doubt that

the hawksbill deserves listing as critically endangered. A few facts are in order. Population declines of the hawksbill have occurred against a backdrop of high levels of international trade in hawksbill shell (tortoiseshell).

Estimates of tortoiseshell imports into Japan from 1970 to 1992 (when Japan stopped trading) exceeded 754 tonnes, representing approximately 712,000 hawksbills. In 1990, hawksbill experts meeting in Nagasaki, Japan, estimated that there were approximately 15,000 hawksbills nesting annually worldwide.

Exploitation of hawksbills for tortoiseshell, meat and eggs continues worldwide. Perhaps the most striking indication of the hawksbill's plight is a recent communication by Colin Limpus of the Queensland Department of Environment and Heritage that researchers in the Solomon Islands are discovering that more than 90% of the nesting hawksbills are first-time breeders (identified by laparoscopy). Thus most females of this long-lived species are not surviving to nest for more than one season.

The only "controversy" about the hawksbill's status is that being created by those who wish to resume international trade in this species. The Cuban hawksbill proposal presented at the CITES meeting in Zimbabwe last June would have allowed

resumption of the harvest of wild hawksbills to supply the luxury tortoiseshell trade between Cuba and Japan. It would also have allowed the sale of all hawksbill shell stockpiled by Cuba since 1992. Both actions would have signalled the resurrection of the destructive tortoiseshell market, which took decades to shut down. Ironically, Japan's need for Cuban stockpile was not even clear — the reported inventory of tortoiseshell in Japan in 1995 was 188 tonnes.

The Cuban hawksbill proposal was defeated, but those who orchestrated its presentation in Zimbabwe are already positioning themselves for their next attempt to reopen the trade in tortoiseshell. Mrosovsky's suggestion that hawksbills should be placed in the 'data deficient' category in the *Red List* is an obfuscating tactic and is appalling in its transparency.

Anne Meylan

Florida Department of
Environmental Protection,
Florida Marine Research Institute,
100 8th Avenue SE,
St Petersburg, Florida 33701-5095, USA
e-mail: meylan_a@harpo.dep.state.fl.us

1. Mrosovsky, N. *Nature* **389**, 436 (1997).
2. Baillie, J. & Groombridge, B. *IUCN Red List of Threatened Animals* (IUCN, Gland, Switzerland, 1996).
3. Groombridge, B. & Luxmoore, R. *The Green Turtle and Hawksbill: World Status, Exploitation and Trade* (Secretariat of CITES, Lausanne, Switzerland, 1989).

Animal experimentation needs dissection

Sir— Ehsan Masood draws attention to the concern, focused by the intention of the UK government to tighten the law on animal experimentation, that research using animals may be curtailed (*Nature* **389**, 896; 1997).

This follows an earlier account of the forthcoming Swiss referendum about the use of transgenic animals (*Nature* **389**, 103; 1997) which expresses the fear that public opinion will lead to a ban, with consequent constraint on biological research. This is pertinent because genetically manipulated animals form an increasing proportion of the total used in experimental procedures.

One problem is that lay perceptions of transgenic animals may be coloured by the organisms that are featured in the popular media — obese, hairless and even luminescent mice — so that the public may get the impression that all genetically engineered organisms are grotesque or bizarre.

We have recently completed a survey of

the views of 778 school and college students about the use of transgenic animals in medical research. Overall, 42% of our sample thought that genetic engineering of laboratory animals for this purpose was wrong, the same proportion that thought using 'ordinary' animals was wrong, suggesting that students find the process of genetically engineering animals for medical research no more or less objectionable than the use of animals in general for such research.

Few students (13%) thought that animal research was acceptable if the animals were caused pain.

It is significant, then, that 19% of the students thought that genetically engineered laboratory animals would be in constant pain, although fewer (12%) of the oldest group, the college students aged 16–19, believed this. On the other hand, 49% overall imagined that use of genetically engineered animals would lead to novel discoveries, and this view increased in the college students to 57%, suggesting that there might be an increasing appreciation of the value of transgenic organisms to research.

If our figures, showing that about half the students objected to animal research, extend to the general population, the outcome of a referendum about animal experimentation might be finely balanced.

There is a need, then, to inform the public in more detail about animal experimentation, including that which involves the production and use of transgenic animals.

Our results suggest that it would be helpful to dissect the issues, and distinguish between pain, distress and discomfort caused to animals.

Ruaraidh Hill
Martin Stanisstreet
Edward Boyes

Environmental Education
Research Unit,
University of Liverpool,
Liverpool L69 3BX, UK
e-mail: martstan@liv.ac.uk

Helen O'Sullivan
Department of Environmental
and Biological Studies,
Liverpool Hope University College,
Liverpool L16 9JD, UK

Nomenclature regulation

Sir—The recent leading article¹ and subsequent correspondence^{2–4} on nomenclature illustrate problems that the scientific community is largely failing to address. A general reluctance to adopt any form of nomenclature regulation will ensure a continuing increase in idiosyncratic names being associated with genes. To try to overcome this problem, the International Society for Plant Molecular Biology, through its Commission of Plant Gene Nomenclature (CPGN), has undertaken to unify gene nomenclature within ‘plants’⁵.

To achieve this goal, all plant nuclear-encoded sequences in the SwissProt and SPTREMBL databases have been grouped by cluster analysis and gene family analysis. This analysis is soon to be expanded to incorporate all eukaryotic organelle sequences. Each gene family has been assigned a unique alphanumeric identifier, whether or not it has an acceptable name. Assigning to each gene family a name that will reflect the function of the proteins within the family is a big task, and to attempt this retrospectively for the many gene families that do not have names will be time-consuming.

The assignment of unique ‘gene family numbers’ has many immediate advantages: it unifies families between species but could also be used in a wider context to unify nomenclature within the eukarya. Gene family numbers can be linked to expressed sequence tags (ESTs) and used to identify the location of the EST on chromosomal maps in place of the myriad of meaningless EST accession numbers. Gene family names would eventually replace the gene family numbers, making maps intelligible.

The CPGN’s database Mendel (<http://jiiio6.jic.bbsrc.ac.uk>) holds some 3,500 sequence records relating to approximately 300 gene families. The current analysis, when complete, will expand the number of gene families to nearer 2,000, adding a further 12,500 records. This information will be released during 1998 and will be reciprocally cross-linked to European Molecular Biology Laboratory sequence records.

In my experience, most people in the scientific community do not know of attempts to standardize nomenclature. Here surely is a role for the journals, which could not only promote the existence of nomenclature databases but could also insist on authors submitting the recognized ‘name’ or, in its absence, the gene family number besides the sequence accession number. This would inevitably force the pace of change in unifying nomenclature.

The CPGN would be pleased to

collaborate with other nomenclature projects to help promote unification of nomenclature within the eukarya.

David Lonsdale

CPGN, Bioinformatics,
John Innes Centre,
Norwich NR4 7UH, UK
e-mail: cpgn@bbsrc.ac.uk

1. *Nature* **389**, 1 (1997).
2. Whyte, B. J. *Nature* **390**, 329 (1997).
3. Povey, S. *et al.* *Nature* **390**, 329 (1997).
4. Puente, L. *et al.* *Nature* **390**, 329 (1997).
5. *Trends Biol. Sci.* **21**, 443 (1996).

...and linguistic diversity

Sir—Lnc Puente *et al.* complain about the large number of gene and protein names that have unclear or offensive pronunciations because they lack vowels (*Nature* **390**, 329; 1997). This is a tunnel-tongued complaint, as there are many beautiful languages that use vowels much less profligately than English; in these languages, such indeterminate names may have a single pronunciation. Indeed, in Czech the name *Lck*, an example that Puente *et al.* give, has a single definitive pronunciation yet no meaning (denigrating or otherwise).

We propose that nomenclature committees embrace the richness of human linguistic diversity by indicating which language should be used to pronounce each gene and protein name. As a first step in that direction, we propose that *Srk*, *Mls* and *Prs* be given their conventional pronunciations.

We prefer linguistic liberation to acronym anarchy.

Daniel E. Weeks

Jeffrey R. O’Connell

Department of Human Genetics,
University of Pittsburgh,
A310 Crabtree Hall,
130 DeSoto Street,
Pittsburgh, Pennsylvania 15261, USA
e-mail: dweeks@watson.hgen.pitt.edu

Zora Schmidtová

Lukešova 30,
Praha 4,
Czech Republic

Disadvantaged women

Sir—Two items in your issue of 4 December, “Female scientists wanted — apply to UK research councils” and “No evidence of sexism in peer review” (*Nature* **390**, 431 & 438; 1997), discuss the lower success rates of female scientists compared to their male colleagues in winning research grants. Both quote an undeniable statistic that a considerably smaller proportion of women apply for the grants in the first

place, and both appear to be baffled by the cause.

In fact the cause is simple. A much smaller proportion of the women who are loosely head-counted as ‘members of the department’ are actually in established positions. Most are supported on short-term contracts and so are debarred from applying for their own research grants from that parent (but often the most appropriate) source.

Thus a large proportion of half the equally qualified adults in a given scientific discipline do not have the opportunities to form groups and attract collaborators. Yet those opportunities are crucial for building up the group expertise and reputations that the scientific world looks for when head-hunting.

If I were to apply for a grant to research this fundamental problem, I would not expect much support.

Elizabeth Griffin

6 London Place,
Oxford OX4 1BD, UK
e-mail: remg@astro.ox.ac.uk

Pragmatism in Latin America

Sir—The coverage of the current bioethics momentum as presented in *Nature* (**389**, 658–663; 1997) is both enlightening and timely. There is, however, no reference to Latin America, which has rather belatedly taken account of bioethics, mostly in the form of ‘solutions’ to dilemmas arising in more mature scientific traditions. But some forms of bioethical reflection and teaching can be found in many Latin American countries, particularly Argentina, Chile, Colombia, Brazil and Mexico.

For example, a master’s degree programme is offered by the University of Chile with the collaboration of the Program on Bioethics of the Pan-American Health Organization established in Santiago de Chile in 1993. There are similar programmes in Argentina, Brazil and Colombia.

Like other countries, those in Latin America are confronted with the task of incorporating into local practices the pragmatism embodied in the bioethical enterprise in the United States and the United Kingdom, as well as of adapting it to their particular cultural idiosyncracies.

Fernando Lolas

(Vice-President,
University of Chile)

Jorge Litvak

(President)
International University Exchange Inc.,
1717 Massachusetts Avenue NW, Suite 802,
Washington, DC 20036, USA

Feathers, filaments and theropod dinosaurs

D. M. Unwin

One of the hottest debates in palaeontology is whether birds evolved from dinosaurs. A study of two exceptionally well-preserved specimens of a theropod dinosaur from China – complete with skin, internal organs and eggs – provides new clues to the origin of feathers.

Ever since John Ostrom resuscitated the idea in the 1970s, palaeontologists have been piling up the evidence in favour of theropod dinosaurs as the ancestors of birds. Recently, the gap between theropods and birds has been narrowed even further. Important avian characters, such as the furcula (wishbone), have been discovered in theropods thought to be close to birds¹. And typical theropod characteristics, such as an enlarged claw on digit two of the foot, have been found in early birds². The final, clinching fact would be the discovery of evidence of feathers — a defining feature of birds — in theropods.

Cue *Sinosauropteryx prima*, the so-called 'feathered' dinosaur from China (Fig. 1). Reports of this sensational discovery first appeared³ in late 1996. Although few scientists have yet seen the fossil material, some are already incorporating *Sinosauropteryx* into models for the origin of feathers and



Figure 1 Reconstruction of *Sinosauropteryx* by Michael W. Skrepnick. Three exceptionally well-preserved specimens of this dinosaur have been recovered from the Early Cretaceous Yixian Formation of China. Two of these are described by Chen *et al.*⁶, who show that *Sinosauropteryx* bore discrete structures that could have been 'proto-feathers'.

bird flight⁴. Still others argue⁵ that the 'feathers' are merely an artefact of preservation. On page 147 of this issue⁶, Chen, Dong and Zhen present the first detailed description of *Sinosauropteryx*, and show that the integument (skin) bore discrete filamentous structures. But are they feathers?

Chen *et al.* describe two almost complete, near-adult individuals, with evidence of soft tissues including the integument, the eyes and possibly some internal organs⁷. They are, without doubt, the best-preserved dinosaur remains yet found. These, along with a third specimen, were recovered by Chinese farmers in Liaoning Province, China (P. J. Currie, personal communication). They were found in beds of the Yixian Formation, part of a thick sequence of lake sediments intercalated with volcanic deposits. During the past six years, these sediments have yielded many superbly preserved fossils. The result is an almost complete Early Cretaceous (145–97.5 million years ago) continental biota, composed of plants, insects, fish, lizards, turtles, pterosaurs, dinosaurs, mammals and

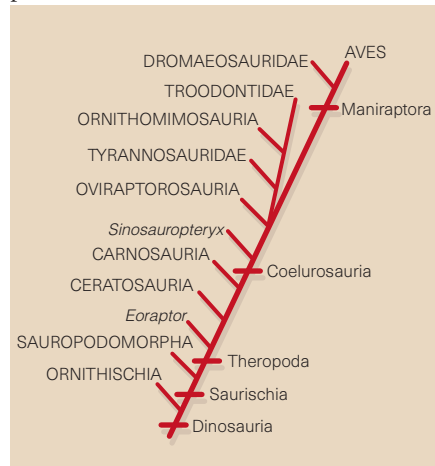


Figure 2 The relationships of *Sinosauropteryx* to other dinosaurs. *Sinosauropteryx* is not as closely related to birds as many other theropods, yet it is the first non-avian theropod that seems to show evidence of feather-like structures. (Modified from ref. 10.)

numerous birds⁸ — the latter often preserved with their plumage intact⁹.

In life, *Sinosauropteryx* was about the size of a large chicken and distinguished by its deep, narrow body, remarkably long tail and rather short, stout forelimbs (Fig. 1). Like its close relative *Compsognathus*, from the Late Jurassic of Europe (163–145 million years ago), *Sinosauropteryx* has a very specialized hand. Its massive first digit bears a large claw that might have served as a killing tool. Last meals provide further evidence of an active predatory lifestyle — a lizard in the gut region of one individual⁶ and a tiny mammal in the third specimen (P. J. Currie, personal communication). The body of the larger individual described by Chen *et al.* also contains another surprise — two small, oval structures which, judging by their shape, size and position, are almost certainly eggs. This is the first reasonably convincing record of this type of association for any dinosaur and, if correctly identified, provides incontrovertible evidence that this individual was female. Moreover, the relatively small size of the eggs and the possibility of paired oviducts⁶ suggest that, unlike modern birds (which have a single oviduct and, in general, smaller clutches of relatively large eggs), theropods had a more reptile-like reproductive system with two oviducts that produced larger numbers of relatively small eggs.

The most striking features of the fossils are the well-preserved remains of the integument, which forms a dark halo above the skull, neck, back, hips and both sides of the tail. Small patches of integument also occur on the skull, and are associated with the forelimbs, rib-cage and legs. The halo is composed of many coarse, sinuous filaments, which are possibly hollow and up to 40 mm long⁶. The filaments seem to be discrete structures and, although often matted and tangled, are not encased in skin or remnants of the decaying dermis, as some have suggested⁵. Branching of the filaments — a construction that is also typical of feathers — has been reported^{6,10}, but the topography of this branching is not yet clear.

This brings us to the critical question: are these structures some kind of 'proto-feather'? Chen *et al.*⁶ and others^{4,10} clearly favour this idea. They draw comparisons with the plumules of modern birds, which are relatively simple structures without barbules or hooklets. But any argument for homology between the feathers of birds and the integumentary structures of *Sinosauropteryx* needs to be supported by more than general similarities in structure and position. High-resolution microscopy and biogeochemical tests might provide some answers, but they will not solve all of the problems. Moreover, if *Sinosauropteryx* bears proto-feathers, we might expect similar (or perhaps even more feather-like)

structures to have been present on at least some of the theropods that are more closely related to birds than is *Sinosauropteryx* (Fig. 2). Exceptionally well-preserved remains of the integument are now known for two of these dinosaurs — the ornithomimosaur *Pelecanimimus*¹¹, and a small, unnamed maniraptoran theropod from Brazil¹². In both cases, however, there is no evidence of the filamentous structures found in *Sinosauropteryx*.

So, it seems that we still do not have absolute proof that some dinosaurs were feathered. Or do we? The Liaoning deposits have yielded three examples of another putative dino-bird intermediate, *Protarchaeopteryx*. According to a preliminary report by Ji and Ji¹³, these individuals have well-preserved evidence of true feathers. While we wait for further details of these Chinese fossils, we might consider the irony of the present situation — nothing for

hundreds of years and then, suddenly, a whole flock of evidence. □

D. M. Unwin is in the Department of Geology, University of Bristol, Wills Memorial Building, Queen's Road, Bristol BS8 1RJ, UK.
e-mail: dave.unwin@uk.ac.bristol

1. Norell, M. A., Mackovicky, P. & Clark, J. M. *Nature* **389**, 447 (1997).
2. Forster, C. A., Chiappe, L. M., Krause, D. W. & Sampson, S. D. *J. Vert. Paleontol.* **16**, (Suppl.) 34A (1996).
3. Hecht, J. *New Scientist* No. 2052, 7 (1996).
4. Padian, K. *J. Vert. Paleontol.* **17**, (Suppl.) 68A (1997).
5. Geist, N. R., Jones, T. D. & Ruben, J. A. *J. Vert. Paleontol.* **17**, (Suppl.) 48A (1997).
6. Chen, P., Dong, Z. & Zhen, S. *Nature* **391**, 147–152 (1998).
7. Ruben, J. A., Jones, T. D., Geist, N. R. & Hillenius, W. J. *Science* **278**, 1267–1270 (1997).
8. Hecht, J. *New Scientist* No. 2094, 32–35 (1997).
9. Hou, L., Martin, L. D., Zhou, Z. & Feduccia, A. *Science* **274**, 1164–1167 (1996).
10. Currie, P. J. in *Encyclopedia of Dinosaurs* (eds Currie, P. J. & Padian, K.) 241 (Academic, San Diego, 1997).
11. Briggs, D. E. G., Wilby, P. R., Pérez-Moreno, B. P., Sanz, J. L. & Fregenal-Martinez, M. *J. Geol. Soc. (Lond.)* **154**, 587–588 (1997).
12. Kellner, A. W. A. *Nature* **379**, 32 (1996).
13. Ji, Q. & Ji, S. A. *Chin. Geol.* **238**, 38–41 (1997).

Cosmology

An unprincipled Universe?

Peter Coles

One of the central tenets of cosmological orthodoxy is the Cosmological Principle, which states that, in a broad-brush sense, the Universe is the same in every place and in every direction. This assumption has enabled cosmologists to obtain relatively simple solutions of Einstein's General Theory of Relativity that describe the dynamical behaviour of the Universe as a whole. These solutions, called the Friedmann models¹, form the basis of the Big Bang theory. But is the Cosmological Principle true? Not according to Sylos-Labini *et al.*², who argue, controversially, that the Universe is not uniform at all, but has a never-ending hierarchical structure in which galaxies group together in clusters which, in turn, group together in superclusters, and so on.

These claims are completely at odds with the Cosmological Principle and therefore with the Friedmann models and the entire Big Bang theory. The central thrust of the work of Sylos-Labini *et al.* is that the statistical methods used by cosmologists to analyse galaxy clustering data are inappropriate because they assume the property of large-scale homogeneity at the outset. If one does not wish to assume this then one must use different methods.

What they do is to assume that the Universe is better described in terms of a fractal set characterized by a fractal dimension D . In a fractal set, the mean number of neighbours of a given galaxy within a volume of radius R is proportional to R^D . If galaxies are distributed uniformly then $D = 3$, as the number of neighbours simply depends on the volume of

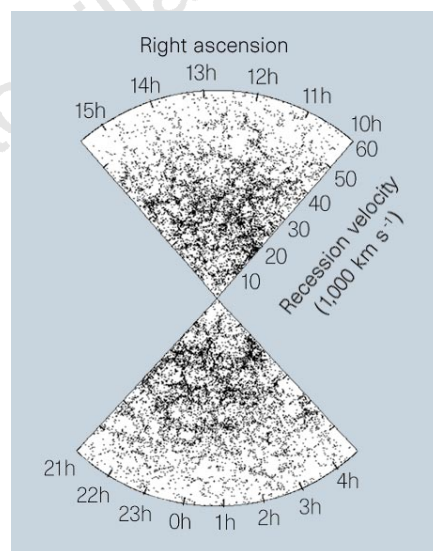


Figure 1 The big picture — the Las Campanas redshift survey, each dot marking a galaxy with a well-determined redshift. The survey maps the Universe out to recession velocities of 60,000 km s⁻¹, corresponding to distances of a few hundred million parsecs. Although no fractal structure on the largest scales is apparent (there are no clear voids or concentrations on the same scale as the whole map), one statistical analysis² finds a fractal dimension of two in this and other surveys, for all scales — conflicting with a basic principle of cosmology.

the sphere and the average number-density of galaxies. A value of $D < 3$ indicates that the galaxies do not fill space in a homogeneous fashion: $D = 1$, for example, would indicate that galaxies were distributed in roughly lin-

ear structures (filaments). Sylos-Labini *et al.* argue that $D = 2$, which suggests a roughly planar (sheet-like) distribution of galaxies.

Most cosmologists would accept that the distribution of galaxies on relatively small scales, up to perhaps a few tens of megaparsecs (Mpc), can indeed be described in terms of a fractal model. This small-scale clustering is expected to be dominated by purely gravitational physics, and gravity has no particular length scale associated with it. But standard theory requires that the fractal dimension should approach the homogeneous value $D = 3$ on large enough scales. According to standard models of cosmological structure formation, this transition should occur on scales of a few hundred Mpc.

The main source of the controversy is that most available three-dimensional maps of galaxy positions are not large enough to encompass the expected transition to homogeneity. Distances must be inferred from redshifts (see box), and it is difficult to construct these maps from redshift surveys, which require spectroscopic studies of large numbers of galaxies.

Sylos-Labini *et al.* have analysed a number of redshift surveys, including the largest so far available, the Las Campanas redshift survey³. They find $D = 2$ for all the data they look at, and argue that there is no transition to homogeneity for scales up to 4,000 Mpc, way beyond the expected turnover. If this were true, it would indeed be bad news for the orthodox among us.

Their results are, however, at variance with the visual appearance of the Las Campanas survey, for example, which certainly seems to display large-scale homogeneity (Fig. 1). Objections to these claims have been lodged by Luigi Guzzo⁴, for instance, who has criticized their handling of the data and has presented independent results that appear to be consistent with a transition to homogeneity. It is also true that Sylos-Labini *et al.* have done their cause no good by basing some conclusions on a heterogeneous compilation of redshifts called the LEDA database⁵, which is not a controlled sample and so is completely unsuitable for this kind of study. Finally, it seems clear that they have substantially overestimated the effective depth of the catalogues they are using. But although their claims remain controversial, the consistency of the results obtained by Sylos-Labini *et al.* is impressive enough to raise doubts about the standard picture.

Mainstream cosmologists are not yet so worried as to abandon the Cosmological Principle. Most are probably quite happy to admit that there is no overwhelming direct evidence in favour of global uniformity from current three-dimensional galaxy catalogues, which are in any case relatively shallow. But this does not mean there is no

Deeply in the red

According to the Hubble law, the observed recession velocity of a galaxy v is proportional to its distance d from the observer: $v = H_0 d$, where the constant of proportionality H_0 is Hubble's constant. For velocities much less

than the velocity of light c , the recession velocity is simply proportional to the fractional change in wavelength of spectral lines emitted by the galaxy due to the Doppler effect. This change, called the redshift, z ($\approx v/c$), is used

to estimate the distance to the galaxy via $d = cz/H_0$. At greater redshifts, where the recession velocities become comparable to c , the form of the law changes, and distance is less than this formula would imply.

P.C.

evidence at all: the near-isotropy of the sky temperature of the cosmic microwave background, the uniformity of the cosmic X-ray background, and the properties of source counts are all difficult to explain unless the Universe is homogeneous on large scales⁶. Moreover, Hubble's law itself is a consequence of large-scale homogeneity: if the Universe were inhomogeneous one would not expect to see a uniform expansion, but an irregular pattern of velocities resulting from large-scale density fluctuations.

But above all, it is the principle of Occam's razor that guides us: in the absence of clear evidence against it, the simplest model compatible with the data is to be preferred. Several observational projects are already under way, including the Sloan Digital Sky Survey and the Anglo-Australian 2DF

survey, that should chart the spatial distribution of galaxies in enough detail to provide an unambiguous answer to the question of large-scale cosmic uniformity. In the meantime, and in the absence of clear evidence against it, the Cosmological Principle remains an essential part of the Big Bang theory. □

Peter Coles is in the Astronomy Unit, Queen Mary and Westfield College, University of London, London E1 4NS, UK.

e-mail: p.coles@qmw.ac.uk

1. Friedmann, A. Z. *Phys.* **10**, 377–386 (1922).
2. Sylos-Labini, F., Montuori, M. & Pietronero, L. *Phys. Rep.* (in the press).
3. Shectman, S. et al. *Astrophys. J.* **470**, 172–188 (1996).
4. Guzzo, L. *New Astron.* **2**, 517–532 (1997).
5. Paturel, G. et al. in *Information and Online Data in Astronomy* (eds Egret, D. & Albrecht, M.) 115 (Kluwer, Dordrecht, 1995).
6. Peebles, P. J. E. *Principles of Physical Cosmology* (Princeton Univ. Press, NJ, 1993).

Structural biology

Synchronized division proteins

Roy Burns

The molecular mechanisms of prokaryotic cell division and eukaryotic mitosis seem to have little in common, yet similarities between the structures of two GTPases published in this week's issue^{1,2} (pages 199 and 203) will reinforce the need for microbiologists and cell biologists to compare notes. Nogales *et al.*¹ report the 3.8-Å image reconstruction of the tubulin α/β heterodimer, which assembles into the microtubules of, for example, eukaryotic mitotic spindles. Löwe and Amos², on the other hand, describe the 2.8-Å crystallographic structure of *Methanococcus jannaschii* FtsZ, which forms a filamentous ring-shaped septum at the division site of prokaryotes.

It has been said that many of the best intelligence analysts are trained as biologists, because they can solve complex problems from fragmentary, and sometimes conflicting, evidence. The 3.8-Å structure¹ of the α/β tubulin dimer is a testament to this skill. It confirms many aspects of the working model built up over the past 30 years from

indirect biochemical analysis, such as the effects due to the reactivity of specific cysteine residues on ligand binding, sequence analysis and site-directed mutagenesis. The structure is, of course, more tangible than any previous model, even though its resolution is insufficient to assign functions to specific residues. But it will greatly influence

studies into the mechanisms of microtubule assembly, microtubule-based motility and the effects of the commercially important anti-mitotic ligands.

Nogales *et al.*¹ have determined the structure of Zn^{2+} -induced sheets of anti-parallel α/β tubulin protofilaments, with GTP bound to the α -subunit, and taxol* and GDP bound to the β -subunit. The apparent simplicity of the microtubule ultrastructure belies the fact that the assembled dimer can adopt a bewildering variety of conformational states. These lead to treadmilling³ (whereby steady-state microtubules preferentially assemble at one end and dissociate at the other) and dynamic instability⁴ (whereby steady-state microtubules show stochastic length changes). Direct measurements of the subunit-dissociation rate constants and the protofilament subunit-subunit periodicity⁵ have established that these states result from assembly-dependent hydrolysis of the β -tubulin GTP, and from differences in the kinetic properties of the two microtubule ends.

Analysis of severed microtubules indicates that there are three conformational states⁶ at each microtubule end. These may correspond to subunits at the two ends with bound GTP, GDP + inorganic phosphate, and GDP. In fact, this may underestimate the complexity, because different states may also be induced by binding the microtubule-associated proteins and other modulators of the assembly kinetics. Indeed, the requirement for this conformational plasticity may partly account for the extraordinarily high conservation of the tubulins — 11 per cent of the β -tubulin residues are completely conserved in a sample of 208 sequences representing almost all phylogenetic orders.

Structural information is going to be needed for each of the conformational states before the dynamic properties of microtubules are fully understood. We also need to know the relevant ligand association and conformational rate constants (many of which are likely to depend on the isotype). The 3.8-Å structure of the tubulin dimer

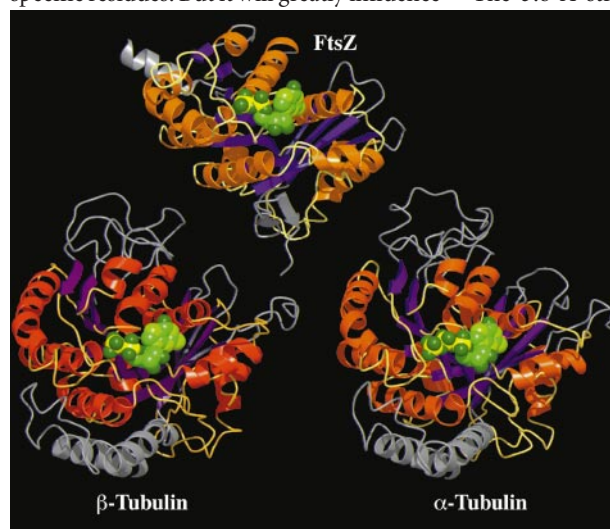


Figure 1 Orientations of FtsZ and the α - and β -tubulins. Nogales *et al.*¹ have solved the structure of the tubulin α/β heterodimer, and Löwe and Amos² have solved the FtsZ structure. The structures are orientated to reveal the nucleotide-binding face. The bound nucleotide is shown in green, equivalent α -helices in orange, and equivalent β -strands in purple. Features that are different in FtsZ and the tubulins — including loops and the carboxy-terminal domain of the tubulins — are shown in grey.

E. NOGALES

*Bristol-Myers Squibb has registered Taxol as a trademark and wishes the scientific community to use the name paclitaxel.

is a veritable *tour de force*, but only one step towards the full story.

Determination of the structure of FtsZ-GDP by Löwe and Amos² may represent another step. FtsZ is thought to be the bacterial homologue of the eukaryotic tubulins⁷, although the sequences are less than 20 per cent identical. But similarities include hydrolysis of the bound GTP during the *in vitro* assembly of FtsZ into filaments; the longitudinal periodicities of these filaments and the microtubule protofilaments; and the assembly of both proteins into sheets and rings or spirals^{8,9}.

The crystallographic structure shows that FtsZ has two main domains. One of these binds a guanine nucleotide at a site that is atypical compared with the more conventional GTPases such as p21^{ras}. The modified Rossman fold of this domain is, however, strikingly similar to that of the tubulin subunits (Fig. 1). Moreover, FtsZ and β -tubulins share many of the residues that are implicated in coordination of the bound nucleotide.

But this structural similarity between the nucleotide-binding domains of FtsZ and the

tubulins presents a paradox. Is it a quaint coincidence, with atypical GTPases having unrelated functions in prokaryotic and eukaryotic cell division? Or have the two protein families evolved from a common ancestor, such that microtubules and FtsZ filaments share some functional properties? If this is the case, FtsZ may provide a way to dissect the conformational complexity of the tubulin subunit — and microbiologists and cell biologists will have a lot to talk to one another about. □

Roy Burns is in the Biophysics Section, Blackett Laboratory, Imperial College of Science, Technology and Medicine, London SW7 2BZ, UK.

e-mail: r.burns@ic.ac.uk

1. Nogales, E., Wolf, S. G. & Downing, K. H. *Nature* **391**, 199–203 (1998).
2. Löwe, J. & Amos, L. A. *Nature* **391**, 203–206 (1998).
3. Margolis, R. L. & Wilson, L. *Cell* **13**, 1–8 (1978).
4. Mitchison, T. J. & Kirschner, M. *Nature* **312**, 237–242 (1984).
5. Hyman, A. A., Chretien, D., Arnal, I. & Wade, R. H. *J. Cell Biol.* **128**, 117–125 (1995).
6. Tran, P. T., Walker, R. A. & Salmon, E. D. *J. Cell Biol.* **138**, 105–117 (1997).
7. Erickson, H. P. *Cell* **80**, 367–370 (1995).
8. Mukherjee, A. & Lutkenhaus, J. *J. Bacteriol.* **176**, 2754–2758 (1994).
9. Erickson, H. P., Taylor, D. W., Taylor, K. A. & Bramhill, D. *Proc. Natl Acad. Sci. USA* **93**, 519–523 (1996).

Particle physics

Neutrino deficit challenges conservation laws

Frank Wilczek

SuperKamiokande is a modern wonder of the world: 50,000 m³ of purified water surrounded by giant phototubes looking out for Cerenkov radiation, the shock waves of light left behind by 'superluminal' particles. (I speak here not of the hypothetical and dubious tachyons, but of ordinary particles boosted to high energy, which outrace light through water even though their speed is less than that of light in a vacuum.) The giant instrumented aquarium has proved to be a superb detector of the rare interactions of neutrinos. At a conference last month*, the SuperKamiokande group reported an anomaly in their observations of neutrino interactions that has considerable implications for fundamental physics (E. Kearns, Boston Univ.). The anomaly is in the relative number of electron-type and muon-type neutrinos found in cosmic rays. The observation greatly strengthens earlier hints in the same direction^{1–5}. It threatens cherished conservation laws, and calls into question the completeness of the Standard Model of particle physics.

Let me first define the matter at stake. A wide range of experimental results can be explained by the postulate that there are conservation laws for three different classes of fundamental particle: the laws of conser-

vation of electron number, muon number and tau number. Known collectively as the laws of lepton-number conservation, these laws make very similar, very simple statements about what sorts of particle inter-

actions do or do not occur.

Electrons, for example, are assigned electron number +1, and their antiparticles, positrons, are assigned electron number –1. Electron neutrinos, ν_e , have electron number +1, and the corresponding antineutrinos $\bar{\nu}_e$ electron number –1. All other particles have electron number 0. In any reaction we have yet observed, the sum of all the particles' electron numbers in the initial state is equal to that in the final state. The laws of muon (μ) and tau (τ) number conservation follow exactly the same pattern.

The idea that electron and muon numbers are separately conserved was postulated⁶ to explain why certain types of particle decay, notably $\mu \rightarrow e\gamma$, in fact never occur (e is an electron; γ is a photon). The 'two-neutrino' experiment⁷, eventually rewarded by a Nobel prize, tested this very idea. It showed that neutrinos produced in the decay of pi mesons, $\pi^- \rightarrow \mu^- \bar{\nu}_\mu$, were able to induce the reaction $\bar{\nu}_\mu p \rightarrow n \mu^+$ but not $\bar{\nu}_\mu p \rightarrow n e^+$ (p is a proton; n is a neutron). The idea has since been tested in a variety of ways, and is now embedded in the Standard Model of particle physics.

The laws of lepton-number conservation greatly resemble, in their formulation and use, the dozens of atom-number conservation laws — one for each type of chemical element — in allowed chemical reactions. These laws are what defeat the dreams of alchemy, and they have served chemistry well; but nuclear physics shows that they are only approximate. SuperKamiokande may be teaching us a similar lesson for the elemental leptons.

So what has SuperKamiokande seen? When energetic cosmic rays strike the atmosphere they typically produce many

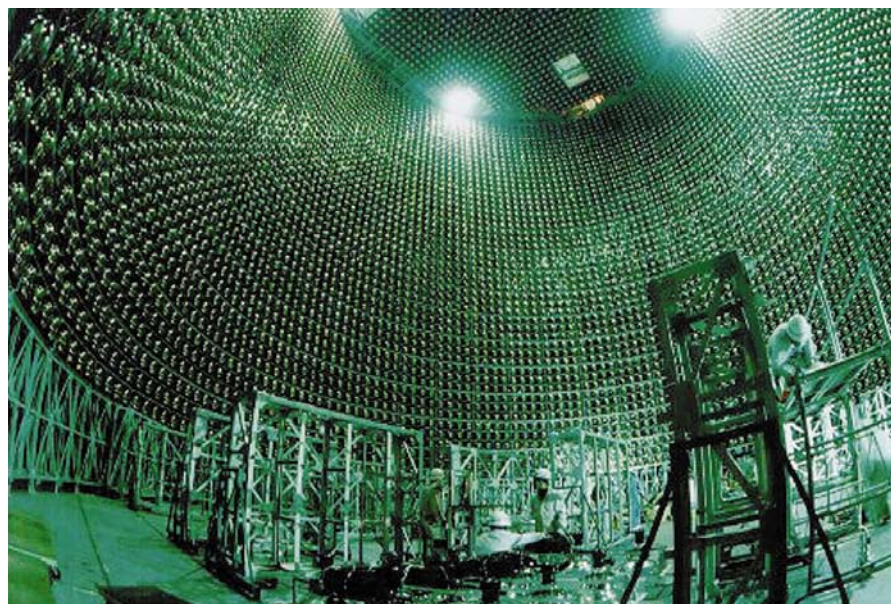


Figure 1 Inside the SuperKamiokande detector. Deep underground, a vast volume of water is surrounded by photomultiplier tubes, which detect flashes from high-energy neutrino interactions. An observed lack of muon-type neutrinos from the atmosphere appears to overturn the laws of lepton-number conservation. (Details of SuperKamiokande can be found at <http://www-sk.icrr.u-tokyo.ac.jp/doc/sk/index.html>.)

* Santa Barbara Neutrino Conference, 2–6 December 1997; <http://www.itp.ucsb.edu>.

charged pi mesons, which in turn decay into states containing muon neutrinos or anti-neutrinos. The muons themselves then decay, usually into a muon neutrino, an electron and an electron antineutrino; and there are several other processes that produce neutrinos of both types.

Although the details are complicated, a great deal is known experimentally about cosmic-ray showers, and essentially all the basic underlying processes have been studied in great detail. As a result, one can estimate the ratio of muon to electron neutrinos with considerable confidence⁸ — it is close to 2:1. But experimenters at SuperKamiokande find far fewer muon neutrinos and anti-neutrinos than were predicted. The observed ratio is about 1:1.

If the theoretical estimates were correctly reckoned, and the experimental results correctly interpreted, what could account for the seeming contradiction? To many physicists, the most interesting possibility is that neutrinos of one type 'oscillate' into neutrinos of another type. Indeed, if many of the mu-type neutrinos making their way from the atmosphere to the SuperKamiokande detector oscillate into tau-type neutrinos, which are more difficult to detect, the observations could be explained. Of course, that would violate the conservation of both muon and tau lepton numbers.

The idea of neutrino oscillations goes back a long way^{9,10}, having been proposed not long after oscillations between different types of K meson were first observed, during the childhood of particle physics. In this view, different kinds of neutrino are like different polarization states of light. In empty space, a given polarization state of light will be maintained as it propagates. But as light passes through a crystal, or any optically active material, one polarization will oscillate into another. Very similar mathematics governs the propagation of neutrinos, so oscillations among the different types would indicate that empty space is 'leptonically active'.

The question of whether neutrino oscillations occur is closely connected to the question of whether neutrinos have mass. Indeed, the oscillations are attributed to what are called off-diagonal masses. Whereas ordinary masses, in quantum mechanics, cause the phase of the wavefunction of a particle to oscillate in time, off-diagonal masses cause not only its phase but also its direction (polarization, for example, or neutrino type) to oscillate.

This connection to mass is central to the deeper theory of neutrino oscillations. For one thing, it allows us to quantify the observed effect. If the SuperKamiokande observations do indicate neutrino oscillations, then the value of the off-diagonal masses responsible is of the order of 10^{-2} electron volts. This is more than six orders of magnitude smaller than the mass of the electron, which itself is by far the

smallest mass appearing in the Standard Model. So the oscillation would only represent a tiny correction to the Standard Model, visible only because a very small oscillation per unit length gets multiplied by the long distance neutrinos can travel between interactions. It does not so much challenge the Standard Model as supplement it.

Most theoretical physicists would welcome such tiny, but non-zero, neutrino masses as an encouraging sign that our ambitious ideas about unification and symmetry beyond the Standard Model may be on the right track. The existence of off-diagonal mass terms for quarks has long been established, and if one attempts to treat quarks and leptons symmetrically — as one must in all modern unification schemes — it is hard to avoid such terms for neutrinos. Furthermore, the most complete and beautiful unification schemes, which view all the quarks and leptons as facets of a single symmetrical entity, require the existence of an additional, heavy, neutral particle. This so-called right-handed neutrino mixes with ordinary (left-handed) neutrinos in a way that induces tiny masses for the latter. These tiny masses are expected, though with considerable uncertainty, to be roughly the size of those claimed by SuperKamiokande.

The new, firm reports of an anomaly in the cosmic-ray neutrinos will be sharpened — or compromised — by additional observations in coming months. Decisive would be observations that the lost muon neutrinos still exist in another form: this might be veri-

fied by studying so-called neutral current interactions, which are largely insensitive to the change.

Finally, we should note that the atmospheric neutrino anomaly occurs in the context of two other claimed neutrino deficits, similar in form but quite different in important details: long-standing claims, from several experiments, of a deficit of electron neutrinos from the Sun¹¹, and recent observations by an accelerator team at Los Alamos¹². These various reports neither confirm nor directly contradict one another — with three different neutrinos and several possible mass terms coming into play, there is room for several distinct oscillation phenomena on different energy and length scales. It seems probable that the separate laws of lepton-number conservation will soon fall. □

Frank Wilczek is in the School of Natural Sciences, Institute for Advanced Study, Olden Lane, Princeton, New Jersey 08540, USA.
e-mail: wilczek@ias.edu

1. Hirata, K. S. *et al. Phys. Lett.* **205**, 416–420 (1988); **280**, 146–152 (1992).
2. Casper, D. *et al. Phys. Rev. Lett.* **66**, 2561–2564 (1991).
3. Becker-Szendy, R. *et al. Phys. Rev. D* **46**, 3720–3724 (1992).
4. Fukuda, Y. *et al. Phys. Lett.* **335**, 237–245 (1994).
5. Allison, W. W. M. *et al. Phys. Lett.* **391**, 491–500 (1997).
6. Feinberg, G. *Phys. Rev.* **110**, 1482–1483 (1958).
7. Danby, G. *et al. Phys. Rev. Lett.* **9**, 36–44 (1962).
8. Gaisser, T. K. *et al. Phys. Rev. D* **54**, 5578–5584 (1996).
9. Pontecorvo, B. *Zh. Eksper. Theoret. Fiz.* **34**, 247–249 (1958); *Sov. Phys. JETP* **7**, 172–173 (1958).
10. Maki, Z. *et al. Prog. Theor. Phys.* **28**, 870–876 (1962).
11. Bahcall, J. & Krastev, P. *Phys. Rev. D* **53**, 4211–4225 (1997).
12. Athanassopoulos, C. *et al. Phys. Rev. Lett.* **77**, 3082–3085 (1996).

Palaeoecology

Did forests survive the cold in a hotspot?

Peter D. Moore

Tropical forests harbour a large proportion of the world's biodiversity, they act as a major reservoir for carbon in the global cycling of that element, and they are under threat. Yet, surprisingly, little is known about their detailed history — particularly in Africa. A group of European palaeoecologists¹ has now sought to close this gap in our knowledge with an intensive study, published in the *Journal of Biogeography*, of the highland region of east-central Africa, running south from the Virunga volcanoes in Uganda, through Rwanda into Burundi (Fig. 1). They conclude that current biodiverse 'hotspots' were not always so species rich.

A number of questions need to be answered regarding the history of the forests in tropical Africa. To what extent were they affected (and possibly fragmented) by the colder climates of the high-latitude glacial episodes? Have they expanded, contracted or moved in the form of intact communities?

Or have they, like their temperate counterparts, changed their composition as individual taxa have responded to climate change according to their own requirements? We also need to know whether the current biodiversity hotspots are located in sites where relict fragments of the forest survived the cold, dry climate of the last glacial events. And to what extent has human activity — even early in prehistory — modified the pattern of forest spread in the Holocene (the past 10,000 years)?

Assembling all of the available information, in 1982, Hamilton² suggested that certain core areas, probably resulting from refuges that had survived the climatic shocks of the Pleistocene (roughly the past two million years), acted as centres for forest spread and recolonization of central Africa during the Holocene. Among these cores were eastern Zaire (including the western fringes of Uganda, Rwanda and Burundi, bordering the Western Rift Valley), the coastal area of

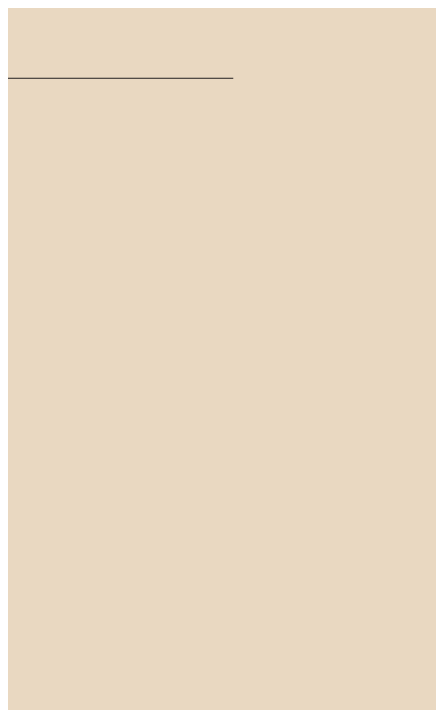


Figure 1 Map of east Africa, showing the swamp locations studied by Jolly *et al.*¹. Their studies of these five locations indicate that the tropical forest in this region came either from a core refugium in the west which survived the climatic shocks of the Pleistocene, or one that was derived from dispersed patches of forest. (Figure adapted from ref. 1.)

eastern Tanzania, and the west coast of Africa in Cameroon/Gabon. Possible minor core areas were also proposed, in northern Mozambique and in Ethiopia. According to Hamilton, the plant and animal species spread from these postulated refugia as they moved into the intervening areas and reassembled the interglacial forest.

To put this hypothesis to the test, palaeoecological information is needed to document patterns of vegetation change in these regions. Moreover, a number of sites must be investigated, to detect differential movements and to avoid the problems that arise from the possible use of an atypical site. In the latest study by Jolly *et al.*¹, five suitable sites were located along the highlands on the eastern side of the Western Rift Valley, running in a north–south transect of about 300 km. All sites were peat-accumulating swamps within the broad-leaved forest zone. Their depths ranged from 7 to 23 metres, and all contained well-preserved pollen. Some of the cores contained material dating from as far back as 18,000 years before present (BP), thus taking the vegetation history record back to the coldest stage of the last glacial episode.

The oldest sediments contained pollen typical of dry ericaceous scrub and grassland. This indicates that the area supported little forest during the height of the cold episode. There is evidence for the elements of montane forest, but the pollen could have



Figure 2 Montane vegetation, to the east of the sites studied by Jolly *et al.*¹.

been transported over some distance in this generally tree-less landscape. Alternatively, there could have been patches of montane forest in particularly favourable locations. Expansion of the montane forest varies in its timing between the sites but, unexpectedly, it seems to occur at the higher-altitude locations before the lower ones (about 11,000 and 10,000 yr BP, respectively). But precise radiocarbon dating is hampered by a radiocarbon plateau, making resolution difficult³.

It is apparent from the pollen diagrams that the composition of the richer forest — which follows after about 10,000 yr BP at all sites — varies between the locations. The genus *Alchornea* (Euphorbiaceae), for example, expands more rapidly and to higher levels at the northern end of the transect than in the south. By contrast, *Macaranga* (also Euphorbiaceae) is more abundant and persistent in the south. Increasing levels of scrub and forest-edge indicators in recent times (from about 2,300 yr BP onwards) suggest forest disturbance by human activity, but the time at which these clearances began varies between sites.

In palaeoecological research, especially in east Africa, single-site studies are the norm. But the regional, multi-site study by Jolly *et al.*¹ provides at least some of the answers to the historical biogeographical questions of the Old World tropics. If there was a core refugium of forest in this region, it certainly did not extend east of the Western Rift Valley, where the uplands were covered by ericaceous scrub during the height of the cold period. Either the refugium was further west in Zaire or, perhaps, we should be considering the survival of much smaller, dispersed patches of forest rather than a major refugial forest block.

When the climate became warmer and wetter in the early Holocene, equatorial forests did not march up the mountains in a phalanx. They seem, instead, to have dropped behind the lines in guerrilla fashion. They then expanded from the favoured locations in which they established themselves, sometimes moving downwards from higher altitudes. Palaeoecological studies of temperate forest⁴ indicate that the forest composition varied from site to site, depending both on local environmental conditions and the expansion patterns of the different



100 YEARS AGO

The personal estate of the late Mr. Alfred Nobel has been valued at 434,093l., of which amount 216,901l. is in England. After a number of personal bequests have been made, Mr. Nobel's will stipulates that the capital of the whole of the remaining realisable property is to form a fund, the interest from which is to be annually divided in five prizes to those who during the preceding year have done most for the benefit of humanity. The interest is to be divided into five equal parts, which are to be awarded in prizes as follows: (1) To him who within the department of natural philosophy has made the most important discovery or invention; (2) to him who has made the most important discovery or improvement in chemistry; (3) to him who has made the most important discovery within the department of physiology or medicine; (4) to him who in literature has produced the most important work in an idealistic direction; and (5) to him who has worked most or best for the fraternisation of the nations and for the abolition or diminution of standing armies, as also for the promotion and propagation of peace. ... The will continues:— "It is my express will that at the distribution of prizes no regard is to be paid to any kind of nationality, so that the most worthy competitor may receive the prize whether he is a Scandinavian or not." From *Nature* 6 January 1898.

50 YEARS AGO

I should be interested to know whether any physiologist who is also a smoker has noticed the following phenomenon, and whether it has been previously reported. Choose a quiet room, fill the mouth with tobacco smoke, and blow gently out through a very small aperture between pursed lips. Careful observation of the fine jet of smoke will show that it has a definite periodicity — that of the heart beat. With a delicately controlled jet it is even possible to make the heart blow a smoke ring at each beat.

From *Nature* 10 January 1948.

Many more extracts like these can be found in *A Bedside Nature: Genius and Eccentricity in Science, 1869–1953*, a 266-page book edited by Walter Gratzer. Contact David Plant. e-mail: subscriptions@nature.com

taxa. The specific constitution of the forest has been dynamic in both time and space, most recently because of human disturbance and clearance for agriculture. The same seems to be true of tropical forest.

Biodiversity hotspots in the tropics are unlikely to be explained simply on the basis of refugium theory. But survival history has undoubtedly played an important part in their current richness. The range of opportunities presented in the Holocene development of ecosystems — both in terms of microhabitat diversity and disturbance

features — must also have contributed to the variety of species combinations that can be achieved and maintained in these areas. Such is the essence of a biodiverse hotspot. □

Peter D. Moore is in the Division of Life Sciences, King's College, Campden Hill Road, London W8 7AH, UK.

1. Jolly, D. et al. *J. Biogeogr.* **24**, 495–512 (1997).
2. Hamilton, A. *Environmental History of East Africa* (Academic, London, 1982).
3. Becker, B., Kromer, B. & Trimborn, P. *Nature* **353**, 647–649 (1991).
4. Jacobson, G. L., Webb, T. & Grimm, E. C. in *The Geology of North America* Vol. K-3 (eds Ruddiman, W. F. & Wright, H. E.) 277–288 (Geol. Soc. Am., New York, 1987).

Extrasolar planets

Back in focus

Geoffrey Marcy

Last year a conflagration of controversy engulfed the first planet ever claimed around a Sun-like star; now it has suddenly been extinguished, leaving the new world unscathed. The original suggestion¹ that a planet exists around the star 51 Pegasi was based on periodic variations in Doppler shifts of the light from the star — presumably caused by wobbling because of the gravitational pull of the planet. But criticism of this interpretation came from David Gray², who reported spectral-line variations that could be caused only by an undulating, pulsating stellar surface, not an orbiting planet. This interpretation called into question the reality of not just the 51 Peg planet, but also the other five detected extrasolar planets, as highlighted by the cover of *Nature* (27 February 1997): “Extrasolar planets: Fading from view”. But now, three papers^{3–5} (including those on page 153 and page 154 of this issue) report that the spectral evidence of pulsations is absent, after all. The empirical case for oscillations has vanished, leaving planets as the only plausible interpretation.

The discovery of the planet around 51 Pegasi by Michel Mayor and Didier Queloz¹ caught the scientific world off guard, not only because it was the first planet found around a Sun-like star, but because the orbital period of 4.2 days implied that it is 20 times closer to its star than the Earth is to the Sun. Such extreme proximity for a Jupiter-mass planet (or any planet) was unanticipated by conven-

tional theories of planet formation.

Scorched by its close star, the planet around 51 Peg was further charred by four phenomenal controversies. First, on the US news programme *Nightline*, David Latham announced the discovery of a second planet orbiting 51 Peg. This detection has proved difficult to confirm. Then David Black wrote⁶ that it was no planet but instead a failed star, a ‘brown dwarf’. His twofold reasoning was that its mass “could well be ten times greater than the 0.6-Jupiter lower limit” and that inward orbital migration by planets was an “extreme mechanism” for transporting Jupiters from the outskirts of a planetary system, where they would form. Xiaopei Pan reported interferometric evidence for a star orbiting 51 Peg, not a planet at all. Finally, David Gray and Artie Hatzes⁷ reported evidence that the spectral lines were changing shape with the 4.2-day period, with only “one chance in about 300 of this alignment occurring by chance”.

This plethora of interpretations represents science at its healthiest. But some of them can be rejected by astrophysical reasoning. For example, the brown dwarf would raise tides on the primary star, and so exert a torque that would spin-up 51 Peg to a 4-day rotation period. Instead, we can tell that the star rotates roughly once every 30 days, based on its weak magnetic field — not much different from that of the Sun. But all of the above threats to the planet interpretation

lose strength in the face of the seven additional planetary companions since discovered around other stars (see Table 1). Their number and variety provide a powerful argument against quirky explanations for any individual detection. Further, the planet discovered around 47 Ursae Majoris⁸ seems quite ‘normal’, having a minimum mass of 2.5 Jupiter masses and orbiting at 2.1 Earth–Sun distances in a nearly circular orbit. The Doppler velocities of 16 Cygni B have the clear signature of an oval orbit, precisely consistent (and predictable) from newtonian physics. By far the simplest explanation for these last two is that they are Jupiter-mass objects at Solar-System-like distances.

The arguments for Gray’s pulsation hypothesis can be subjected to a *posteriori* review. Some people have already expressed dissatisfaction that the original paper was not thorough in its statistical treatment (citing no error estimates, for example), and used less than equivocal language about the spectral-line variations (“the chance of their being caused by a planet is vanishingly small”). But these critics, including the present author, may occasionally forget that competition and human emotion have always provided fuel for the vigorous pursuit of alternative theories. It was right that the planet interpretation should not go unchallenged.

The scrutiny of 51 Peg is reminiscent of that surrounding the erroneous discovery by Andrew Lyne and collaborators of the “First planet outside our Solar System”, orbiting a pulsar⁹. Lyne subsequently realized that the signal was simply an artefact of the elliptical orbit of the Earth. His honest retraction highlighted the scientific process at its finest. As Lyne did, in withdrawing that apparent detection, Gray⁴ has quickly resolved this controversy by reporting new, more plentiful data that do not show pulsations. In the end we should be impressed by the exquisite care with which Mayor and Queloz examined every alternative interpretation right from the start — the absurd hypothesis of a Jupiter-mass companion in a 4.2-day orbit faced its most severe inquisition from the discoverers themselves.

Apparently, a question posed by Epicurus and Aristotle in 400 BC now has an empirical answer: other worlds — large ones at least — are common. □

Geoffrey Marcy is in the Department of Astronomy, University of California at Berkeley, Berkeley, California 94720, and San Francisco State University, San Francisco, California 94132, USA. e-mail: gmarcy@etoile.berkeley.edu

Table 1 Planets around other stars

Star	Distance from Earth (parsecs)	Mass $\times \sin(\text{inclination}^*)$ (Jupiter masses)	Semi-major axis (AU)	Period (days)	Eccentricity
51 Pegasi	15.4	0.44	0.051	4.2308	0.01
Upsilon Andromedae	16.5	0.63	0.053	4.621	0.03
Rho1 55 Cancri	13.4	0.85	0.12	14.656	0.03
Rho Corona Borealis	16.7	1.1	0.23	39.6	0.05
16 Cygni B	~22	1.74	1.70	802.8	0.68
47 Ursae Majoris	14.1	2.42	2.08	1,093	0.09
Tau Bootis	~15	3.64	0.042	3.3126	0.0
70 Virginis	18.1	6.84	0.47	116.7	0.40

*The orbital inclination to our line of sight is unknown.

1. Mayor, M. & Queloz, D. *Nature* **378**, 355–359 (1995).
2. Gray, D. F. *Nature* **385**, 795–796 (1997).
3. Brown, T. et al. *Astrophys. J.* (in the press).
4. Gray, D. F. *Nature* **391**, 153–154 (1998).
5. Hatzes, A. P., Cochran, W. D. & Bakker, E. J. *Nature* **391**, 154–156 (1998).
6. Black, D. *Sky Telescope* **92**, 20–24 (1996).
7. Gray, D. F. & Hatzes, A. *Astrophys. J.* **490**, 412–424 (1997).
8. Butler, R. P. & Marcy, G. *Astrophys. J.* **464**, L153–L154 (1996).
9. Bailes, M., Lyne, A. G. & Shemar, S. L. *Nature* **352**, 311–313 (1991).

Petroleum geochemistry

Sulphur greases the wheels

Jeffrey S. Seewald

What, in detail, are the fundamental chemical reactions that generate petroleum? Quite simply, we don't know. Such information is essential in developing predictive models for locating and exploiting deposits of oil and natural gas, but has remained elusive due to the complexity of the sedimentary systems in which petroleum forms. In his paper on page 164 of this issue¹, Lewan takes a large step in the right direction — he presents a new hypothesis, based on laboratory experiments, that gives organically derived sulphur species a pivotal role in regulating the rate of petroleum generation.

Lewan proposes that sulphur radicals initiate the breaking of C–C bonds within kerogen (insoluble sedimentary organic matter) to produce the many hydrocarbon fragments that constitute oil, the sulphur radicals being derived from the thermal decomposition of kerogen itself. This view compares with the conventional wisdom that holds that the decomposition of kerogen

to form oil is strictly a thermally driven, kinetic process, dominated by the strength of molecular bonding^{2,3}. Thus, in the past, the inherent weakness of C–S and S–S bonds has been seen as being responsible for the faster generation of petroleum from kerogens that are enriched in sulphur.

In contrast, Lewan suggests that although structural weaknesses are responsible for the generation of sulphur radicals, it is the activity of these radicals, once formed, that facilitates cleavage of C–C bonds during petroleum generation, not the formation process itself. If this idea holds up, it is highly significant — the implication is that more rapid petroleum generation from sulphur-rich kerogen stems from the immediate geochemical environment, and not from a structural characteristic of the kerogen macromolecule.

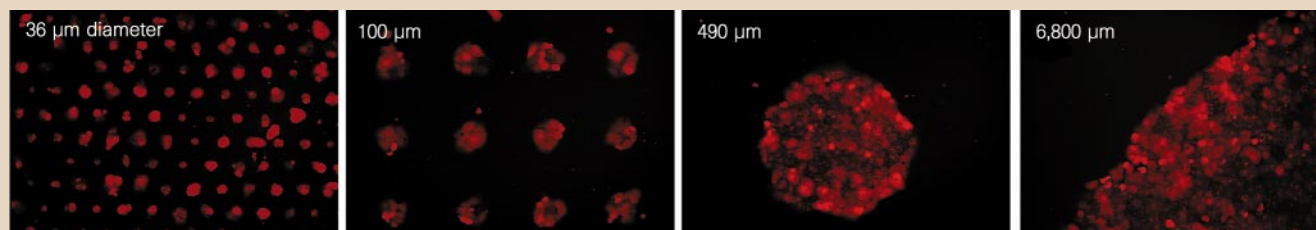
In support of his hypothesis, Lewan provides compelling evidence from laboratory experiments in which 1-phenyldodecane, a model compound intended to represent

kerogen, was heated in the presence of varying concentrations of sulphur radicals. The results show a strong positive correlation between the rate of 1-phenyldodecane cracking and the abundance of sulphur radicals. Lewan concludes that because sulphur-rich kerogen can release more sulphur radicals during maturation, initiation of cracking reactions involving C–C bonds is enhanced and petroleum generation occurs at lower thermal stress. An important feature of Lewan's experiments is that they were conducted in a closed system, allowing secondary reactions to take place between initial reactants and products.

Those who employ relatively simple models of petroleum occurrence, which consider only irreversible reactions, time and temperature^{2,3}, may be disturbed by the idea that the geochemical environment surrounding organic compounds influences chemical reactions. Many of these models rely on kinetic data obtained from open-system pyrolysis, where organic matter is heated under dry conditions at ambient pressure in a stream of inert carrier gas. Rapid removal of reaction products minimizes secondary reactions. As Lewan emphasizes, these experiments do not replicate the geochemical environment

Artificial tissues

Liver spots



The homily that the human body is more than the sum of its parts becomes all too apparent when you have to assemble its components from scratch. The liver, for instance, is a highly structured organ, and its function is not reproduced by a culture of hepatocytes (liver cells) alone. Liver-like function (producing enzymes, filtering out toxins and so on) is retained when hepatocytes are instead co-cultured with tissue-forming fibroblast cells — but a random co-culture is still a poor substitute for real liver tissue. At a meeting last month*, Mehmet Toner (Harvard Med. School) explained how a marriage of cell biology with microfabrication can do better.

The artificial liver — a device for supporting liver-failure patients awaiting a transplant — is not a new idea. The Minnesota-based Regenerex company, for

example, already produces a machine in which hepatocytes sustained in hollow fibres by a perfusion of nutrients do their job on the patient's blood. But such devices are prone to cell leakage, cell death and inefficiency.

A synthetic liver-like tissue is a more attractive option. Unfortunately, fibroblasts and hepatocytes cultured together do not automatically adopt the arrangement found in the liver; and because the two then communicate poorly, the hepatocytes are inefficient. So Toner and colleagues have used patterned substrates to grow the two cell types in more controlled geometric relationships.

They use the standard photolithographic techniques of microelectronic technology to imprint the surface of a borosilicate wafer with patterned films of collagen, which promotes cell adhesion. When hepatocytes

are cultured on the surface, they adhere only to the collagen-coated regions, and the rest of the culture can be washed away (as shown in these images of stained hepatocytes). Fibroblasts can then be introduced to the regions of bare surface, producing an intimate mixture of the two cell types in a periodic pattern.

In this way, an arbitrary ratio of cell types can be prepared on the substrate, so adjusting the ratio to its physiological value is straightforward. Patches of hepatocytes perform best (as measured by their level of albumin or urea secretion) when they are small, because they can only communicate well with the surrounding fibroblast cells up to a certain distance from the interface — a fact revealed by stains selective to albumin or urea production. So the best choice of pattern element size, shape and number density can be deduced and engineered to order in this artificial tissue.

Philip Ball

* Fall Meeting of the Materials Research Society, Boston, 1–5 December 1997.

surrounding kerogen in sedimentary basins, where reactants are confined at high pressure (several hundred bars). So it could well be that reaction mechanisms and model predictions based on open-system pyrolysis experiments have little relevance to processes occurring in natural systems.

Yet the implications of the geochemical environment having a large part in regulating thermal maturation of sedimentary organic matter go well beyond the design of laboratory experiments. Kerogen is just one component of sedimentary rocks, which may contain a grab bag of inorganic minerals and aqueous fluids of extremely diverse composition. Theoretical, field and experimental studies have delivered increasing evidence to support the contention of some researchers, including myself, that the inorganic components of sediments are highly reactive and may participate directly in organic geochemical processes. These components may act as catalysts or regulate key chemical variables such as redox state and pH, which in turn will control thermodynamic drives that influence the direction and extent of chemical reactions, both organic and inorganic⁴⁻⁶. In addition, aqueous pore fluids may represent a reactive source of hydrogen and oxygen for the creation of hydrocarbons and oxygenated organic compounds^{6,7}.

Sulphur is ubiquitous in natural systems, and occurs in its elemental form and an entire range of oxidation states in minerals and aqueous pore fluids. It is intriguing to consider the effect these sulphur sources may have during the thermal maturation of sedimentary organic matter because the presence of various inorganic sulphur species greatly enhances the rates and specificity of numerous organic reactions. These effects are not limited to free-radical reaction mechanisms. For example, the decomposition of aqueous alkanes and alkenes to form methane and carbon dioxide proceeds more rapidly in the presence of aqueous sulphur species in intermediate oxidation states⁸. Other studies have shown increased rates for organic-matter oxidation by aqueous sulphate in the presence of elemental sulphur^{9,10}.

Given the high reactivity of sulphur species with organic molecules, such species are also likely to influence the stability of petroleum after its expulsion from its source rock. In particular, elemental sulphur produced during thermochemical sulphate reduction in petroleum reservoirs may be an effective source of sulphur radicals; these radicals could enhance the cracking of oil to produce natural gas by the same mechanism that Lewan has proposed for the generation of oil from kerogen. The presence of water and aqueous sulphur species in sediment pore fluids may also represent a medium through which the degradation of oil can proceed by mechanisms other than thermal cracking.

Petroleum persists in some geological

environments to much higher temperatures than existing models predict¹¹, and a probable explanation is the switching on and off of specific reaction mechanisms in response to the availability of organic and inorganic reactants and catalysts. As petroleum resources become increasingly scarce, the need for predictive models that include explicit provision for the chemical reactions involved will necessarily increase. Well-designed laboratory experiments, such as those presented by Lewan, represent a powerful means to identify and quantify the relevant geochemical processes to include in these models. □

Jeffrey S. Seewald is in the Department of Marine Chemistry and Geochemistry, Woods Hole

Insect physiology

Caterpillars have lungs

Peter J. Mill

For the past 350 years¹, it has been the perceived wisdom that air is distributed to insect tissues through a series of branched tubes called tracheae and tracheoles, the main branches of the tracheae opening to the outside world at a series of segmentally arranged openings called spiracles. But a paper in the latest issue of the *Journal of Insect Physiology*² gives us cause for a rethink. After a painstaking examination of over 2,000 electron micrographs, Michael Locke has challenged the universality of this 'rule' and concluded that many caterpillars possess lungs, designed to provide blood cells (haemocytes) with oxygen.

The tracheal cuticle (with its underlying epithelium) is continuous with that of the insect's body surface, and the flow of air in and out of the spiracles is often controlled by valves^{3,4}. The tracheae themselves are continuous with the cuticular tubes of the intracellular tracheoles, which form branches that ramify throughout the tissues. The respiratory needs of the tissues are met by diffusion within the lumen of this tubular system, and by the exchange of oxygen and carbon dioxide across the thin walls of the tracheoles^{5,6}. This mechanism applies to most tissues.

However, the pair of spiracles closest to the rear of many caterpillars — including the Brazilian skipper butterfly, *Calpododes ethlius* — are much larger than the others. Moreover, one of the tracheae from each of these spiracles gives rise to a branched tuft of very fine and thin-walled tracheal branches and tracheoles (Fig. 2, overleaf). The tracheae that give rise to these tufts are termed aerating tracheae to distinguish them from the other, conducting, tracheae. The tip of the tuft is anchored in the heart muscle, but the tracheoles do not ramify amongst the cells as in other tissues. Furthermore, much of the tuft is free in the haemolymph, and the tuft

Oceanographic Institution, Woods Hole, Massachusetts 02543, USA.
e-mail: jseewald@whoi.edu

1. Lewan, M. D. *Nature* **391**, 164–166 (1998).
2. Hunt, J. M. *Petroleum Geochemistry and Geology* (Freeman, San Francisco, 1996).
3. Tissot, B. P. & Welte, D. H. *Petroleum Formation and Occurrence* (Springer, Berlin, 1984).
4. Helgeson, H. C., Knox, A. M., Owens, C. E. & Shock, E. L. *Geochim. Cosmochim. Acta* **57**, 3295–3339 (1993).
5. Mango, F. D., Hightower, J. W. & James, A. T. *Nature* **368**, 536–538 (1994).
6. Seewald, J. S. *Nature* **370**, 285–287 (1994).
7. Lewan, M. D. *Geochim. Cosmochim. Acta* **61**, 3691–3723 (1997).
8. Seewald, J. S. *Mater. Res. Soc. Symp. Proc.* **432**, 317–331 (1997).
9. Toland, A. G. J. *Am. Chem. Soc.* **82**, 1911–1916 (1960).
10. Goldhaber, M. B. & Orr, W. L. in *Geochemical Transformations of Sedimentary Sulfur* (eds Vairavamurthy, M. A. & Schoonen, M. A. A.) 412–425 (Am. Chem. Soc., Washington, DC, 1995).
11. Price, L. C. *Geochim. Cosmochim. Acta* **57**, 3261–3280 (1993).

moves as the heart contracts and relaxes².

Amazingly, although the existence and movement of these tufts has been known for more than 100 years⁷, only now has the significance of this been recognized². Locke has found that the haemolymph becomes oxygenated as it passes through the tufts, which waft it towards the openings into the heart. Even more significantly, he has shown that the tufts provide way-stations for haemocytes to replenish their oxygen supplies.

Insect haemolymph contains over seven million haemocytes per millilitre in the *C. ethlius* caterpillar², and there are several types of haemocyte. These engulf or encapsulate foreign material (including invading organisms)⁸, and may also be important in storing and distributing nutrients and/or maintaining sugar and amino-acid levels^{8,9}. In most insects, the haemolymph is not involved in oxygen transport, and haemo-



Figure 1 Alice confronts one caterpillar that clearly has lungs.



Figure 2 The tracheal system in the last three segments of a live caterpillar. Respiration in insects occurs by gaseous diffusion along the lumen of tracheal tubes that branch repeatedly. Blood cells lack their own permanent tracheae, but alight intermittently on a pair of specialized, thin-walled tufts that act like lungs.

cytes were thought to be able to operate in the absence of oxygen. But this has now been called into question².

Many of the haemocytes are loosely attached to tissue surfaces⁸ and, in *C. ethlius*, oxygen deficiency leads to a 60 per cent rise in the number in circulation to nearly 12 million per millilitre of haemolymph. Locke studied the structure of one type of haemocyte — the granulocyte. In a normal, well-oxygenated environment, granulocytes have irregular outlines, with processes (filopodia) on their surface. Their Golgi complexes are well developed and their secretory vesicles contain tubular structures. Under conditions of oxygen starvation, however, they become rounded and lose their filopodia, their Golgi complexes become smaller, and few (if any) tubular structures are seen in their vesicles². Although there are always granulocytes trapped temporarily in the branches of the tracheal tufts, these become more abundant under oxygen starvation (Fig. 3a). Moreover, the granulocytes in the tufts have the characteristics of those in a well-oxygenated environment (Fig. 3b). This is a strong indication that the granulocytes are, indeed, coming to the tracheal tufts for oxygen².

But the story does not finish here. The tracheal branch that gives rise to the tufts also passes backwards into a semi-isolated compartment at the rear end of the animal. Here, again, it becomes thin-walled, giving rise to tracheoles that end freely in the haemolymph and often run back along the sides of the tracheae to form knots. This compartment is called the tokus, and it acts as a 'lung' for the haemocytes. Haemolymph that has passed over the tufts periodically enters the tokus

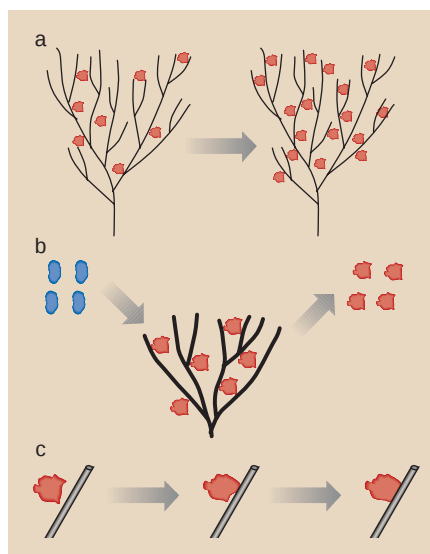


Figure 3 The branched tufts of tracheae and tracheoles that provide blood cells (haemocytes) with oxygen in caterpillars of the butterfly *Calpodethlius*. a, The number of haemocytes (red) in a tuft increases when a caterpillar is subjected to oxygen starvation. b, Oxygen-starved granulocytes (a type of haemocyte; blue) entering a tuft resume the characteristics of those in a well-oxygenated environment (red). c, In the tokus — a 'lung'-like compartment — the haemocytes become closely apposed to the thin-walled tracheae and tracheoles.

through openings alongside the tracheae. The granulocytes (and the other haemocytes) become associated with the free tracheal branches and tracheoles, but in a closer and longer-lasting fashion than in the tufts. The surface of the haemocyte that touches a tracheole becomes flattened to form a large contact area (Fig. 3c), increasing oxygenation of the haemocyte. Because the tokus is close to the heart, the aerated haemolymph can enter it immediately for recirculation².

Locke's work is particularly exciting because it is not just a one-off phenomenon — indeed, the tufts and the tokus lung tracheae are found in caterpillars from all 13 families of lepidopterans that have been studied so far². And Locke has clearly shown that, "although as a rule insect tracheae go to tissues.... haemocytes go to tracheae"². □

Peter J. Mill is in the School of Biology, University of Leeds, Leeds LS2 9JT, UK.

e-mail: p.j.mill@leeds.ac.uk

- Swammerdam, J. *The Book of Nature; or, The History of Insects* (John Hill, London, 1637–1680).
- Locke, M. J. *Insect Physiol.* **44**, 1–20 (1998).
- McArthur, J. M. J. *Exp. Zool.* **53**, 117–128 (1929).
- Miller, P. L. J. *Exp. Biol.* **37**, 237–263 (1960).
- Krogh, A. *Pflüg. Archiv. Ges. Physiol. Menschen Tiere* **179**, 95–120 (1920).
- Wigglesworth, V. B. in *Advances in Insect Physiology* (eds Berridge, M. J., Treherne, J. E. & Wigglesworth, V. B.) 85–148 (Academic, London, 1983).
- King, H. S. *Psyche* **III**, 322–324 (1880).
- Gillott, C. *Entomology* 2nd edn (Plenum, New York, 1995).
- Crossley, A. C. *Adv. Insect Physiol.* **11**, 117–221 (1975).

Daedalus

Speed learning

The geomagnetic field is said to influence psychiatric hospital admissions — though whether it acts on the patients or the diagnostic skill of the admissions staff is not clear. And numerous permanent-magnet health gadgets are sold to those neurotics still at large. How can they work?

Daedalus reckons they stir the fluid axoplasm in the nerves. When a nerve fires, the axoplasm ions move radially. A magnetic field would deflect positive ions in one direction and negative ones in the other, stirring the fluid. And last week Daedalus decided that axoplasmic stirring greatly aids neural adjustment and learning. It speeds the delivery of material needed to update the information held by a synapse.

This makes excellent evolutionary sense. Even in the weak geomagnetic field, the more often a nerve fires, the faster its axoplasm will be stirred. The more it has to learn, the more its learning will be encouraged. Magnetic health gadgets may well help the nerves to adapt to local trouble, reducing pain or weakness.

So Daedalus is improving their design. His new resonant magnetic stimulator has a field reversing at 400 Hz. It thus reverses once during the typical firing-time of a brain nerve. This will stir the nerve powerfully: ions displaced during their outward movement will be displaced even further during their return flow. The axial flow of the axoplasm will be far more vigorous, speeding the nerve's adaptation.

The big use of the tuned stimulator, however, will not be in pain relief, but in education. Our vastly extended childish learning period is no longer adequate. Education fills our entire childhood, and much of young adulthood as well. It is, of course, largely a protection racket, designed to deny young people jobs they could learn perfectly well without the pointless paper 'qualifications' it forces them to waste years acquiring.

But with Daedalus's 400-Hz magnetic hat, they could outwit the system. They wouldn't think any faster; but their brains would be updated twice as fast. They would learn at twice the rate, and finish their education by the age of 10 or so, before puberty steps in. They would no longer turn into truculent teenagers, sexually mature adults resentful at their continued dependency; they could find work, marry, and become useful citizens, as they did in pre-technological society. And the malign youth culture which now exploits their plight would wither mercifully away.

David Jones

Turrell's tunnelling

Turner delights his audiences with his magical depictions of nature on canvas. James Turrell is using nature itself as the canvas — by digging into a volcanic crater to create a new kind of observatory.

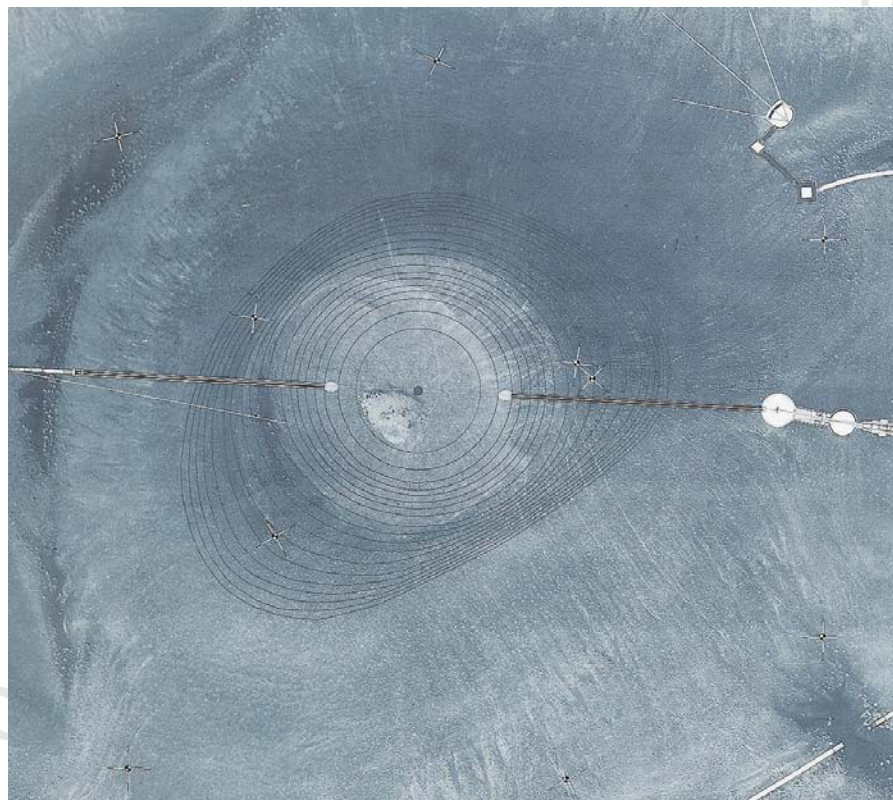
Martin Kemp

Turner, the magician of light and colour who was the dominant figure in British Romantic painting in the first half of the nineteenth century, dealt less with conventional boxes of pictorial space than with “the wide concave of the circumambient air”. His phrase should be understood in the context of his intuition that “the building of Nature” was “too colossal for the intellectual capacity, its height to measure or its depth to fathom — the Universe and Infinity”.

Turner's visual resources were limited to the traditional medium of paint on flat surfaces — if ‘limited’ is the right word for a painter whose floating visions of space, light and colour suggest more in the way of “Infinity” than seems possible.

By contrast, the American artist James Turrell is creating spatial structures on a huge scale to sculpt the passage of natural light and even to reshape the sky itself — in a Turnerian quest to stretch out to infinite reaches of time and space.

In 1975, Turrell searched for a suitable site for the realization of his vast vision of a new “Skyspace”. An aerial survey of the western United States drew his attention to Roden Crater, a cinder cone at the east of the San Francisco volcanic plateau. The project started there in 1979 and is due to be completed early in the next millennium. It involves the excavation of geometrical spaces that will serve as exploratories for our perceptions, as vessels for the enclosure of solar, lunar and stellar illumination, as devices that inscribe time, like a sundial or gnomon, and as a compound observatory in which precision, wonder and imagination are united.



James Turrell's Roden Crater — Negative Site Plan, 1992 (below) and detail of crater (above).

At the centre is the crater transformed into a bowl with an even rim, 3,000 feet in diameter. For the observer low in the bowl, the vault of the heavens is transformed. Subjectively perceived as a shallow dome — a phenomenon much debated in psychology — the sky appears to be anchored circumferentially to the crater's rim. Around the crater, chambers are being tunnelled out in such a way that the passage of light — night into day — will paint the interiors with a succession of shaping shadows and luminous colours.

The fumarole, or secondary vent, will be

the focus for chambers oriented towards the four compass points, joined to a series of linked interiors to provide varying spatial perceptions, and a tunnel 1,035 feet long leading into an oval sky chamber as an atrium for the main bowl. The viewer's subsequent progress to the top of the rim culminates in the reattachment of the dome of the sky to the far horizon of the land.

Turrell's project is presented through models, paintings, drawings, prints and photo-works. Its resonances are astonishingly rich. They range from the prehistoric observatories of stone circles and ancient temples, through the great astronomical castle and garden of Tycho Brahe on the Danish island of Hven, and the camera obscuras of Jesuit magicians in the seventeenth century, to the huge dishes of modern observatories.

At the project's heart is a vision worthy of Turner: “looking at light in a Skyspace is akin to a worldless thought” — and such a thought has boundless potential. □

Martin Kemp is in the Department of the History of Art, University of Oxford, 35 Beaumont Street, Oxford OX1 2PG, UK.



Fish with fingers?

Fingers and toes were long thought to be novelties associated with the invasion of land by tetrapods. The recent identification of a variety of aquatic specializations in some early tetrapods has provoked a debate on whether digits arose in primarily terrestrial or aquatic animals^{1,2}. We recently discovered a pectoral fin of a lobe-finned fish (Fig. 1a, b) that is remarkably similar to tetrapod limbs. This discovery reveals that major tetrapod novelties are also seen in the paddles of some closely related fish and therefore need not have arisen to meet the demands of a terrestrial existence.

We discovered the new fin along a roadside in north central Pennsylvania (Catskill Formation; Late Devonian; ~370 Myr BP). Features of the shoulder girdle and fin rays clearly identify it as belonging to a rhizodontid sarcopterygian³. The specimen compares favourably with *Sauripterus*, known from a single specimen discovered in 1840, in having a broadly flattened radius. The quality of preservation of the new fin permits comparisons with limbs that were impossible until now.

The closest relatives of tetrapods, 'osteolepiform' and elpistostegid sarcopterygians⁴, have fins designed on a simple bifurcate pattern and do not have structures that can be readily compared with digits (Fig. 1c). In contrast, the new fin contains an array of eight distally facing and jointed preaxial radials that are superficially similar in number and configuration to the digits of early tetrapods (Fig. 1d). In addition, six of these radials articulate with homologues of the carpus (intermedium and ulnare) at a common proximo-distal level. The radials of the new fin differ from tetrapod digits in that they are flattened and encased by stiff unjointed dermal fin rays (Fig. 1a, b).

Taken together, these characteristics suggest that the digit-like structures were not the primary load-bearing elements of the distal portion of the rhizodont appendage.

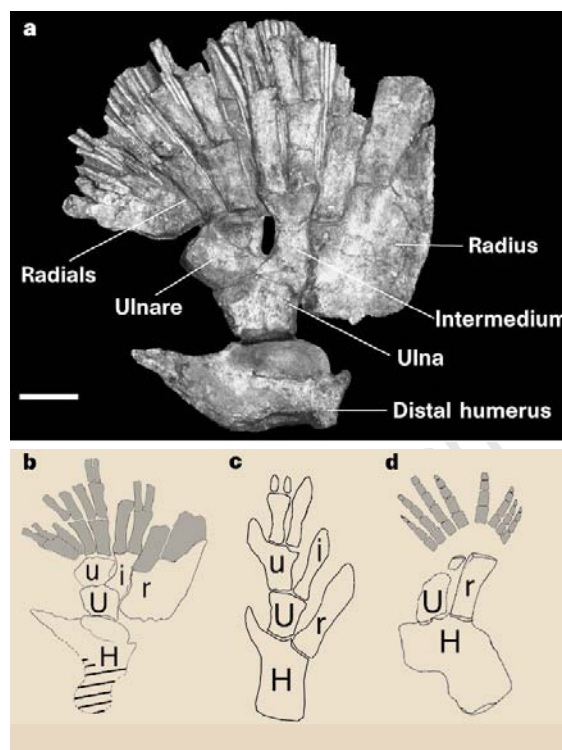


Figure 1 Comparison of fish and tetrapod pectoral appendages. **a**, Ventral view of newly discovered right pectoral fin of a rhizodontid sarcopterygian fish (ANSP 20581); proximal humerus not exposed. Scale bar, 2 cm. Note (1) the array of jointed and branched preaxial radials; (2) common proximo-distal termination of ulnare and intermedium; (3) lack of postaxial process on the ulnare. **b**, Line drawing of ANSP 20581, proximal humerus reconstructed. **c**, Pectoral fin of the osteolepiform sarcopterygian *Eusthenopteron*. **d**, Pectoral limb of the Late Devonian tetrapod *Acanthostega*. Shaded elements emphasize the similarity in pattern of fin radials and digits. H, Humerus; U, ulna; r, radius; u, ulnare; i, intermedium; **b**, **c**, **d** are not to scale.

Phylogenetic comparisons with other sarcopterygians are complicated by missing data but current hypotheses place rhizodontids just outside a group containing 'osteolepiformes', elpistostegalids and tetrapods^{4,5}. This suggests two possibilities: fingers are either primitive to stem tetrapods, or digit-like structures evolved independently in a closely related group of Devonian fish. Either phylogenetic interpretation forces us to question the use of digits as a key innovation associated with the origin of tetrapods.

This discovery reveals that some Devonian fish acquired a truly mosaic fin skeleton, possessing both an extensive and limb-like endoskeleton and elaborate dermal fin rays. The presence of digit-like structures in the paddle of an aquatic fish

suggests that digits could have evolved for reasons other than bearing weight during terrestrial locomotion.

Edward B. Daeschler

Neil Shubin

*Vertebrate Biology,
Academy of Natural Sciences of Philadelphia,
1900 Benjamin Franklin Parkway, Philadelphia,
Pennsylvania 19103, USA
and Department of Biology,
University of Pennsylvania, Philadelphia,
Pennsylvania 19104, USA*

1. Clack, J. A. & Coates, M. I. *Bull. Mus. Natl Hist. Nat., Paris* **17**, 359–372 (1995).
2. Ahlberg, P. E. & Milner, A. R. *Nature* **368**, 507–514 (1994).
3. Long, J. A. *J. Vertebr. Paleontol.* **9** (1), 1–17 (1989).
4. Cloutier, R. & Ahlberg, P. E. *Interrelationships of Fishes*, 445–479 (Academic, New York, 1996).
5. Schultze, H.-P. *Origins of Higher Groups of Tetrapods*, 29–67 (Comstock, Ithaca, 1991).

Takhtajania perrieri rediscovered

Spectacular finds of early Cretaceous fossil flowers during the past decade have fuelled a resurgence of research into the origin of flowering plants¹, the 'abominable mystery' that so captivated Darwin. Now *Takhtajania perrieri* (Capuron) Baranova & J.-F. Leroy, the only extant Africa/Madagascar representative of the family Winteraceae, has been rediscovered in northeastern Madagascar, 85 years after its

original finding. Scientists will thus once again have a 'living fossil' to study and place in the context of both extinct and other extant primitive angiosperms.

At the suggestion of Peter H. Raven, beginning in 1974 with the late Alwyn Gentry, Missouri Botanical Garden botanists and their Malagasy colleagues searched in vain at the Manongarivo Special Reserve in northwestern Madagascar, where French botanist Henri Perrier de la Bâthie had collected the only specimen in 1909. In 1994, the Malagasy plant collector Fanja Rasoavimbahoaka, performing a botanical inventory as part of a joint pro-

gramme of the Missouri Botanical Garden and the World Wide Fund for Nature, collected a flowering tree in the Anjahanaribe-Sud Special Reserve southwest of Andapa. One of us (G.E.S.) identified the specimens as *Takhtajania*. The new locality is 150 km to the east of the original collection. A large population of over 250 adults has been found at the new site, where new research material has already been collected: wood samples, leaves in silica gel for DNA extraction, flowers and fruit in spirit.

The Winteraceae, with about 65 species which exhibit a relict Gondwanan distribution, are thought to be among the oldest

families of flowering plants (along with Chloranthaceae) on the basis of fossil pollen types assigned to modern families. The Cretaceous fossil pollen genus *Walkeripollis* from the late Barremian – early Aptian of Gabon and the late Aptian – early Albian of Israel is winteraceous^{2,3}. Winteraceae persisted in continental Africa until at least the lower Miocene in the southwestern Cape, represented there by two different pollen types associated with modern-day *Bubbia* and *Tasmannia* (sometimes considered *Drimys* section *Tasmannia*), both now restricted to Australo-malesia⁴. The pollen of *Takhtajania*, which is the largest reported in the family, has been compared to that of *Drimys* sensu stricto, which is distributed only in the New World⁵.

Recent molecular phylogenetic studies utilizing internal transcribed spacer (ITS) sequences of ribosomal DNA aimed at resolving genetic relationships within Winteraceae⁶ largely corroborated proposed relationships based on morphological features⁷, but also clearly illustrate how widely divergent *Tasmannia* is from *Drimys*. Preliminary analysis of the *Takhtajania* ITS sequence indicates a basal position for the genus within Winteraceae (Elizabeth Zimmer, personal communication).

Takhtajania perrieri was originally described in 1963 by the French forest botanist René Capuron in the genus *Bubbia*, known from Australia, New Caledonia and the Lord Howe Islands⁸. On the basis of several anomalous features unique within Winteraceae, including a putative bicarpellate ovary and a primitive arrangement of undifferentiated cells surrounding the stomata, Leroy and Baranova in 1978 created a new genus and subfamily to accommodate the species, naming it in honour of the Russian plant systematist Armen Takhtajan⁹. Recent anatomical studies utilizing the type material support the hypothesis of a compound bicarpellate ovary¹⁰. In St Louis in June 1997, while celebrating his 87th birthday and the publication of his latest synthesis on the classification of flowering plants¹¹, Takhtajan was presented with a specimen of *Takhtajania* for the Komarov Botanical Institute in St Petersburg.

George E. Schatz,
Porter P. Lowry II
Missouri Botanical Garden,
P.O. Box 299, St. Louis,
Missouri 63166, USA
e-mail: schatz@mobot.mobot.org
Annick Ramisamihantanirina

Missouri Botanical Garden, B.P. 3391, 101
Antananarivo, Madagascar

- 1273–1275 (1983).
3. Doyle, J. A., Hotton, C. L. & Ward, J. V. *Am. J. Bot.* **77**, 1544–1557; 1558–1568 (1990).
4. Coetzee, J. A. & Muller, J. *Ann. Missouri Bot. Gard.* **71**, 1088–1099 (1984).
5. Pragowski, J. in *World Pollen and Spore Flora* (ed. Nilsson, S. T.) **8**, 1–36 (Almqvist & Wiksell, Stockholm, 1979).
6. Suh, Y., Thien, L. B., Reeve, H. E. & Zimmer, E. A. *Am. J. Bot.* **80**, 1042–1055 (1993).
7. Vink, W. *Taxon* **37**, 691–698 (1988).
8. Capuron, R. *Adansonia* (new series) **3**, 373–378 (1963).
9. Leroy, J.-F. *Adansonia* (new series) **17**, 385–395 (1978).
10. Derooin, T. & Leroy, J.-F. *C.R. Acad. Sci., sér. 3* **316**, 725–729 (1993).
11. Takhtajan, A. *Diversity and Classification of Flowering Plants* (Columbia Univ. Press, New York, 1997).

Changed perceptions in Braille readers

The mature mammalian nervous system has a striking capacity for plastic remodeling in response to environmental changes, but little is known about the perceptual and behavioural relevance of this phenomenon. Using magnetic source imaging we show that the cortical somatosensory representation of the fingers is topographically disordered in blind Braille readers who use three fingers on both hands to read. In addition, they frequently misperceive which of these fingers is being touched. This correlation is suggestive of a functional role for cortical reorganization in the perceptual experience of these individuals.

The somatosensory cortical representation of fingers expands when the amount of sensory input arriving in that part of the brain is increased^{1–6}. Indeed, we have previously used magnetic source imaging to

show that the somatosensory cortical representation of the left hand of string players, which engages in the complex and demanding task of fingering the strings, is expanded¹. To test the hypothesis that fusion of cortical representations is caused by increased simultaneous stimulation^{7–9}, we tested the somatotopic representations of four blind Braille readers who used three fingers (digits 2–4) of both hands simultaneously for reading, and were instructors of the method typically engaged in this practice for several hours a day. We also tested six Braille readers who employed one finger for reading, and five sighted non-Braille-reading subjects. Magnetic source imaging was used to determine the centre of cortical magnetic responsivity to light tactile stimulation of the finger tips (left and right D1, D2 and D5) and right and left lower lip¹.

There was a substantial enlargement of the hand representation in the three-finger Braille readers compared with the other two groups (three-finger readers, 14 mm; one-finger readers, 8 mm; sighted, 7 mm; $F_{2,12} = 16.6$, $P < 0.001$). The three-finger readers also differed from sighted subjects in the topographical arrangements of the finger representations along the postcentral gyrus (Fig. 1). In all sighted subjects the expected homuncular pattern was observed; D1 being inferior (lateral), D2 being more medial, and D5 being the most superior (medial). In each of the three-finger Braille readers, this pattern was changed in one or both hemispheres. Only one of the one-finger Braille readers' cortical topography of finger representations was disordered. These are significant differences ($P < 0.005$, Fisher's exact test, 2-tailed),

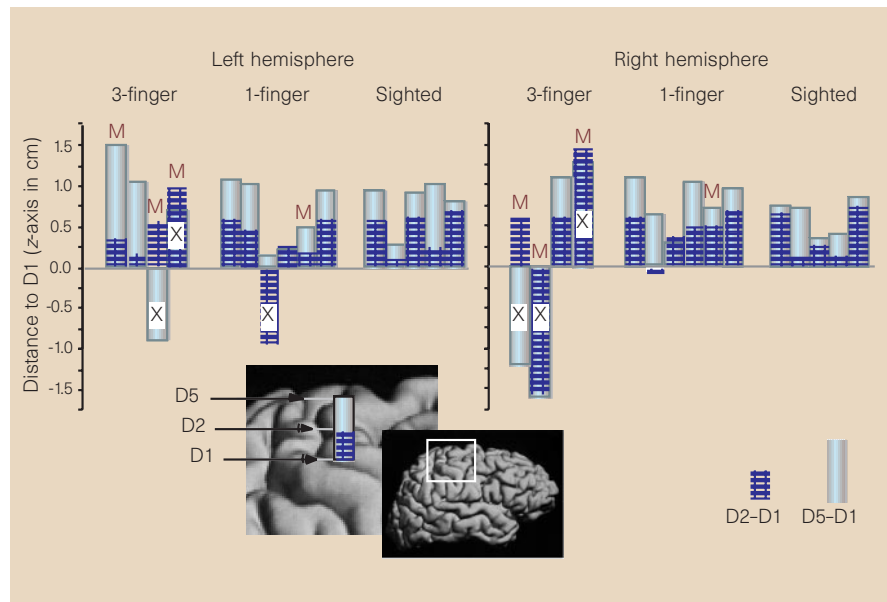


Figure 1 Distances between centres of cortical responsivity: D1, thumb; D2, index finger; D5, little finger. Distances are displayed with reference to the D1 location, which is graphed at the zero-line. The inset shows the centre of responsivity for D1, D2 and D5 superimposed on the brain of a sighted subject. X, Hemispheres in which the homuncular organization of the cortical representations of the fingers was disordered; M, cortical representation of hands in which the location of tactile stimulation was mislocalized.

1. Friis, E. M., Pedersen, K. R. & Crane, P. R. *Pl. Syst. Evol.* (Suppl.), **8**, 31–49 (1994).
2. Walker, J. W., Brenner, G. J. & Walker, A. G. *Science* **220**,

with the three-finger readers differing significantly from the one-finger readers ($P < 0.02$) and sighted controls ($P < 0.008$), whereas the latter groups were not significantly different from one another.

Each of the three-finger readers had a strong tendency to misidentify which finger was being touched during tactile sensory threshold determination, although there was no difficulty in determining that one of the fingers had been touched. In contrast, none of the sighted subjects and only one of the one-finger readers reported such difficulties. Again these are significant differences (for the three groups, $P < 0.005$; three-finger against one-finger, $P < 0.02$; three-finger against sighted, $P < 0.008$; Fisher's exact test, 2-tailed). Sensory threshold testing was also conducted in five other sighted individuals, none of whom had difficulty in localizing tactile stimuli.

A striking feature of the data was the coincidence between the digital topographic disorder in the cerebral cortex and the mislocalization of tactile stimulation of the fingers. This relationship was significant when the data from all individuals were pooled (Fisher's exact test, $P = 0.017$). For the three-finger Braille readers, the hand opposite each of the hemispheres in which there was disorder exhibited a corresponding tactile mislocalization.

'Smearing' of the digital cortical representation could be adaptive for Braille readers who use three fingers in that it serves to fuse input transmitted over different fingers, so that the incoming information can be processed as a whole. Thus, use-dependent cortical reorganization can be associated with functionally relevant changes in the perceptual capacity of the individual.

Annette Sterr, Matthias M. Müller,

Thomas Elbert, Brigitte Rockstroh

Department of Psychology, University of Konstanz,
Fach D25, D-78457 Konstanz, Germany

Christo Pantev

Institute of Experimental Audiology,
University of Münster, D-48129 Münster, Germany

Edward Taub

Department of Psychology,
University of Alabama at Birmingham,
Birmingham, Alabama 35294, USA

Ancient trees in Amazonia

The ages of tropical rain forest trees provide critical information for understanding the dynamics of tree populations, determining historical patterns of disturbance, developing sustainable forestry practices and calculating carbon cycling rates. Nevertheless, the ecological life history of most tropical trees is unknown and even the ages of the largest trees remain to be determined. Tree ages are typically measured by counting annual rings, but in tropical forests rings can be non-existent, annual or irregular¹. In the absence of annual rings, ¹⁴C dating is the only way to determine the age of a tree directly. We have ¹⁴C-dated twenty large, emergent trees from a central Amazon rain forest and find that, contrary to conventional views, trees in these forests can be more than 1,400 years old.

There is debate over how ages are distributed in rain forest trees, although there are almost no direct measurements. Typically, ages are based on extrapolations from growth or mortality rates compiled from permanent plots where observation intervals are short (usually less than 15 years), compared with the longevity of most trees. Because rates vary within a species, and over time within an individual, age estimates are subject to error. Using permanent plot data for 21 canopy species, a study² in Costa Rica calculated that the greatest time required to grow from a diameter of 10 cm to the maximum was 440 years. Another study in Costa Rica estimated that, on the basis of maximum and median growth rates for five emergent species, it would take between 90 and 600 years to reach 100 cm in diameter³.

Supporting younger age estimates, it has been suggested that only trees that exhibit optimum growth rates emerge from the canopy, and that most suppressed individuals are destined to die⁴. In contrast, on the basis of mortality rates, a study in Panama suggested that some trees can live for more than 1,000 years⁵. Our understanding of tropical tree age demographics can be advanced by measuring ages directly, yet the only previously ¹⁴C-dated tree in the Amazon is a 500-year-old⁶ Brazil nut tree (*Bertholletia excelsa*) that is 225 cm in diameter.

To determine the distribution of ages between some large emergent trees in the central Amazon, we ¹⁴C-dated twenty trees from thirteen species harvested in a 80,000-hectare logging operation near the city of Manaus, Brazil. Some trees were very old, with calibrated⁷ ages (± 80 years) ranging from 200 to 1,400 years. Tree size was significantly correlated with age, but most variability was accounted for by other fac-

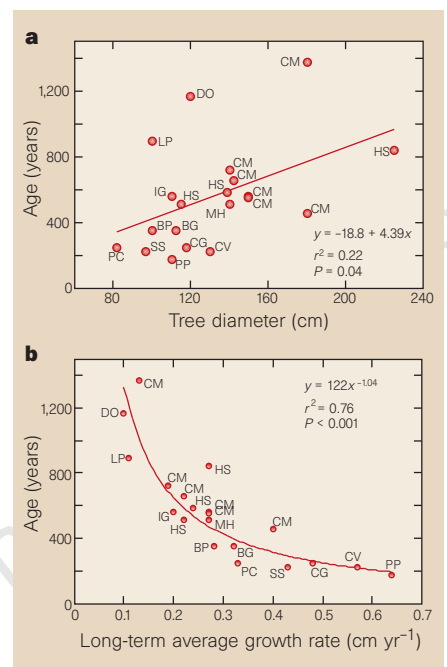


Figure 1 Radiocarbon dates for central Amazon trees. Measurements were made at the Center for Accelerator Mass Spectrometry at Lawrence Livermore National Laboratory. Long-term average growth rates were calculated as diameter divided by age. Although significant, size is not a reliable predictor of age (a). For large trees, growth rate is probably a much better predictor of age (b). Species are: BG, *Bagassa guianensis*; BP, *Brosimum parinarioides*; CM, *Cariniana micrantha*; CG, *Caryocar glabrum*; CV, *Caryocar villosum*; DO, *Dipteryx odorata*; HS, *Hymenolobium* spp.; IG, *Iryanthera grandis*; LP, *Lecythis poiteaui*; MH, *Manilkara huberi*; PP, *Parkia pendula*; PC, *Peltogyne catingae*; SS, *Sclerolobium* spp.

tors (Fig. 1a). Long-term average growth rate was highly correlated with age and the relationship was nonlinear (Fig. 1b).

Average growth rates varied from 1.0 to 6.4 mm yr⁻¹, and within one species alone (*Cariniana micrantha*) rates varied from 1.3 to 4.0 mm yr⁻¹. This suggests that trees can emerge from the canopy by rapid growth, presumably in gaps; by protracted slow growth; or by some combination of both, and the strategy is probably highly species-dependent. At a larger scale, ages for the oldest trees in a region set the minimum time since the last catastrophic disturbance. There is evidence that extensive drought, and perhaps widespread fires, linked to large El Niño events occurred in the Amazon basin 1,500, 1000, 700 and 450 years ago⁸. If the central Amazon was affected, some trees must have survived a number of these events.

Age data also provide important information for foresters. If commercial-sized trees are many centuries old, developing sustainable forest management practices that result in limited forest structural changes will require either huge tracts of forests or very long harvest cycles. Finally,

1. Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B. & Taub, E. *Science* **220**, 21–23 (1995).
2. Elbert, T., Sterr, A., Flor, H., Rockstroh, B., Knecht, S. & Taub, E. *Exp. Brain Res.* **117**, 161–164 (1997).
3. Recanzone, G. H., Jenkins, W. M. & Merzenich, M. M. *J. Neurophysiol.* **67**, 1015–1030 (1992).
4. Recanzone, G. H., Merzenich, M. M. & Jenkins, W. M. *J. Neurophysiol.* **67**, 1057–1070 (1992).
5. Recanzone, G. H., Merzenich, M. M., Jenkins, W. M., Grajski, A. & Dinse, H. R. *J. Neurophysiol.* **67**, 1031–1056 (1992).
6. Jenkins, W. M., Merzenich, M. M., Ochs, M. T., Allard, T. & Guic-Robles, E. *J. Neurophysiol.* **63**, 82–104 (1990).
7. Allard, T., Clark, S. A., Jenkins, W. M. & Merzenich, M. M. *J. Neurophysiol.* **66**, 1048–1058 (1991).
8. Wang, X., Merzenich, M. M., Sameshima, K. & Jenkins, W. M. *Nature* **378**, 71–75 (1995).
9. Byl, N. N., Merzenich, M. M. & Jenkins, W. M. *Neurology* **47**, 508–520 (1996).

ancient trees have implications for the residence time of carbon in wood. From permanent plot data we know that about 50% of the above-ground biomass is contained in less than the largest 10% of the trees. Thus, although they are few, the largest trees represent a sizeable component of the forest's carbon budget, and the associated carbon can be very old. These results provide a first look at how ages are distributed among emergent trees of the central Amazon, and underscore the importance of age demographics in the ecological structure and function of these forests.

Jeffrey Q. Chambers

Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, California 93106, USA
and Health and Ecological Assessment Division, Lawrence Livermore National Laboratory, Livermore, California 94550, USA
e-mail: chambers@lifesci.ucsb.edu

Niro Higuchi

Instituto Nacional de Pesquisas da Amazônia (INPA), CP 478, 69.011 Manaus-Amazonas, Brasil

Joshua P. Schimel

Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, California 93106, USA

1. Fahh, A., Burley, J., Longman, K. & Maraun, A. in *Age and Growth Rate of Tropical Trees: New Directions for Research* (eds Bormann, F. & Berlyn, G.) 31–54 (Yale Univ. Press, New Haven, 1981).
2. Lieberman, D. & Lieberman M. J. *Trop. Ecol.* **3**, 347–358 (1987).
3. Clark, D. A. & Clark, D. B. *Ecol. Monogr.* **62**, 315–344 (1992).
4. Ashton, P. in *Age and Growth Rate of Tropical Trees: New Directions for Research* (eds Bormann, F. & Berlyn, G.) 3–6 (Yale Univ. Press, New Haven, 1981).
5. Condit, R., Hubbell, S. P. & Foster, R. B. *Ecol. Monogr.* **65**, 419–439 (1995).
6. Camargo, P. B., Salomão, R. P., Trumbore, S. & Martinelli, L. A. *Sci. Agricola* **51**, 389–391 (1994).
7. Stuiver, M. & Becker, B. *Radiocarbon* **28**, 863–910 (1986).
8. Meggers, B. *Clim. Change* **28**, 321–338 (1994).

Stratification in poured granular heaps

The stratification of poured granular mixtures into layers according to particle size has long been identified as an important mechanism by which such materials segregate^{1,2}. The implications of this effect for the process industries have also been discussed^{3,4}. Makse *et al.*⁵ suggest that stratification takes place only when there is a marked difference in the shape of large and small grains; specifically, when the angle of repose of the coarse phase is much greater than that of the fines. Our experiments show that stratification can occur in the absence of such heterogeneity of particle shapes within the mixture.

We conclude that stratification of granular mixtures is a multidimensional issue. It is first necessary to establish a specific domain of interest within material param-

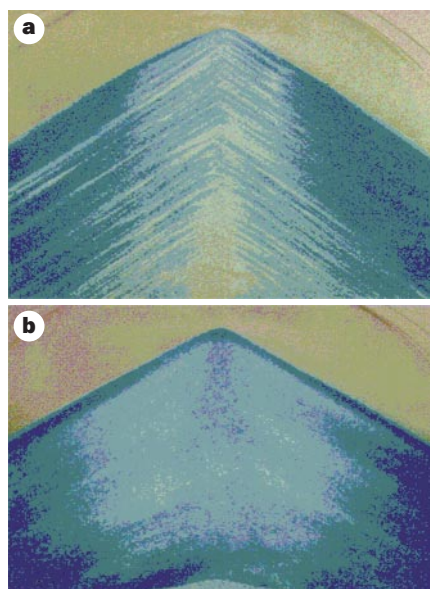


Figure 1 **a**, Stratified heap formed in slow feeding process; **b**, well-mixed heap formed in fast feeding process.

ter space. We consider non-percolating granular flows, where the size and density ratios between coarse and fine phases are sufficiently modest to prevent percolation of fines through the coarse phase from being the primary mechanism for segregation. Percolation is known to dominate if the size ratio is sufficiently high⁶. Under these conditions the fine phase drains through the coarse and is found near the bottom of the assembly.

In a series of experiments, we poured a 20-kg mixture of borax pentahydrate granules into a quasi-two-dimensional hemispherical structure (Fig. 1a, b). The structure was 1 m in diameter, with two plates of Perspex separated by a gap of about 9 cm. The granulate had a continuous distribution of sizes from about 0.1 to 1.4 mm. The granules greater than 0.85 mm in diameter were regarded as the coarse phase and were coloured blue; the fine grains were white. The size ratio of the two phases, in terms of d_{50} measurements, was approximately 2:1 (0.51 and 1.00 mm for fine and coarse phases respectively), and the particle density of each phase was almost identical. The angles of repose of the two phases were 26.9° for the fines and 27.8° for the coarse, so there was negligible difference between the two. The results show that the rate at which material is poured into the structure is a crucial parameter. Figure 1a shows the heap formed in a slow filling process, with a fill time of 45 min (average fill rate 0.007 kg s⁻¹). There is marked stratification of the material into layers of fine and coarse. In contrast, Fig. 1b shows the heap formed with a fill time of 25 s (average fill rate 0.8 kg s⁻¹). The fine and coarse phases are relatively well mixed, with no stratification.

Stratification thus depends strongly on the fill rate, which parametrizes the overall impact (kinetic energy) of the incident feed stream on the growing heap. This effect has previously been identified for bulk segregation^{7,8}. The controlling mechanism is of 'capture' of incident grains by the growing assembly. If the overall impact is high incident grains can embed themselves within the growing assembly, so their freedom of movement is greatly restricted (provided that the mixture is non-percolating). If capture is relatively inefficient, the grains are relatively free to move, so coarser grains will roll over finer grains and stratified layers will be formed.

Only the efficiency of capture is affected by the shapes of the grains in the two phases. For example, the probability of a large grain finding a geometrically stable position in which to embed itself, within an assembly of mostly smaller, near-spherical grains, is relatively low. Thus it is not, as has been previously suggested, a necessary condition for stratification that the angle of repose of the coarse grains should exceed that of the fines. Only the range of fill rates for stratification is affected by the particle shape profile within the mixture. If the angle of repose of the coarse grains greatly exceeds that of the fines, stratification will occur over a wide range of fill rates. However, if the feed rate is sufficiently slow, stratification can occur even if the angle of repose of the fines exceeds that of the coarse grains. In a further experiment, not shown here, we increased the overall angle of repose of the fine phase to 32° by adding a higher proportion of very fine material. We again saw stratification, but only under very slow filling conditions.

We are continuing to examine the relative importances of fill rate, size ratio and particle shape for stratification within poured granular mixtures.

John Baxter, Ugur Tüzün

Department of Chemical and Process Engineering, University of Surrey, Guildford, Surrey GU2 5XH, UK

David Heyes

Department of Chemistry, University of Surrey, Guildford, Surrey GU2 5XH, UK

Igan Hayati, Per Fredlund

Technology Department, Borax Europe Ltd, 170 Priestley Road, Guildford, Surrey GU2 5RQ, UK

1. Williams, J. C. *Univ. Sheffield Fuel Soc. J.* **14**, 29–34 (1963).
2. Williams, J. C. *Powder Technol.* **2**, 13–20 (1968).
3. Carson, J. W., Royal, T. A. & Goodwill, D. J. *Bulk Solids Handling* **6**, 139–144 (1986).
4. Carson, J. W., Royal, T. A. & Goodwill, D. J. *Chem. Eng. Prog.* **92**(11), 72–80 (1996).
5. Makse, H. A., Havlin, S., King, P. R. & Stanley, H. E. *Nature* **386**, 379–382 (1997).
6. Tüzün, U. & Artega, P. *Chem. Eng. Sci.* **47**, 1619–1634 (1992).
7. Drahn, J. A. & Bridgwater, J. *Powder Technol.* **36**, 39–53 (1983).
8. Standish, N., Yu, A. B. & He, Q. L. *Powder Technol.* **68**, 187–193 (1991).

The long walk northwards

Nansen: The Explorer as Hero

by Roland Huntford

Duckworth: 1997. Pp.598. £25

Derek Fordham

"Strange there is always sadness on departure. It is as if I cannot after all bear to leave this bleak waste of ice, glacier, cold and toil." Fridtjof Nansen's sentiments when he returned in 1896 from his journey across the Arctic Ocean reveal something of the restless melancholic nature of this outstanding explorer and statesman, the subject of Roland Huntford's latest work in a series that has so far included Roald Amundsen, Robert Scott and Ernest Shackleton.

Nansen had already made the first crossing of Greenland in 1888, and began his rise from being a moderately successful career scientist, researching into the central nervous system, to becoming one of the great innovators who changed the face of polar exploration. He was at his best when he elected to strike across Greenland from the virtually unknown east coast, thereby eliminating any temptation to turn back, with "Death or the west coast" as his motto.

Suspended between snow and sky in the unknown white heart of Greenland, his expedition marked a revolution in the concept of exploration. The Victorian romantic ideal of enduring suffering as accomplishment was superseded by what Nansen called merely "a ski trip". In Norway he was borne along on a wave of national fervour, and within 18 months he had married, undertaken several lecture tours, written *The First Crossing of Greenland*, and begun to plan his next expedition.

The North Pole was the next objective, and Nansen stood a reigning concept on its head by planning to deliberately freeze a specially designed boat, the *Fram*, into the ice north of Siberia and then be carried by the transpolar drift over the North Pole. He felt he would thus "take note of the forces of nature and work with them".

This was a dangerous and lengthy project, and Huntford meticulously shows how Nansen's turbulent, unpredictable nature seriously threatened harmony on board *Fram*, but was balanced by the steadfastness of Otto Sverdrup, his companion from Greenland and the ship's master. They were characters defined by the distinction between personality and leadership. Nansen was the man with drive, Sverdrup with the talent for command.

When it became clear that *Fram* would not drift over the North Pole, Nansen and Hjalmar Johansen left the ice-locked vessel at latitude 84° 14', having established a new farthest-north record, and began a hair-



Nansen and his wife Eva, pictured soon after their marriage, not in the Arctic but in a studio pose.

raising, hazardous journey of nearly 400 miles across the shifting, breaking ice of the Arctic Ocean towards Franz Josef Land, the position of which had not yet been precisely established.

With no idea where they were when they found land, Nansen and Johansen built a windowless stone hut in which a hibernatory third polar winter was spent, still sharing the same sleeping-bag and still not on first-name terms. In the spring of 1896 some basic instinct drove them south rather than to the

west where they thought Spitsbergen and salvation lay.

Dogs were heard, a man was seen, and Nansen met Frederick Jackson, leader of the North Pole expedition sponsored by the English newspaper magnate Albert Harmsworth, in an Arctic version of the encounter between Stanley and Livingstone. Against all odds, Nansen had turned defeat into victory. Since leaving *Fram*, he and Johansen had travelled 700 miles across the polar ice and had met Jackson just before



In the summer of 1894, Nansen's vessel *Fram* is still heading north, destined to fall short of the pole.

attempting the crossing to Spitsbergen which they might well not have survived.

Regarding his farthest-north record as under threat, Nansen was anxious to capitalize on it as quickly as possible, and lecture tour followed lecture tour. His absences, following his three years in the Arctic, did nothing to re-establish relations with his wife.

Then in 1900 came the news he had dreaded, that the Italian explorer Amadeo had reached 86° 33' north. Nansen's farthest-north record had stood for only five years. He suffered a further blow when Amundsen reached the South Pole — the objective Amundsen had set himself — for, despite the acclaim Nansen received, he had not reached his own objective, the North Pole. At the age of 40 he needed a role.

He left the world of his Arctic triumphs and, in between various romantic affairs, entered a tumbling kaleidoscope of diplomacy and statesmanship of ever-increasing complexity. Norwegian independence, the League of Nations food and refugee relief and repatriation of prisoners of war all fell into the hands of this great Viking statesman. Despite all the international machinations associated with this work, and the accompanying encounters with Lenin, Trotsky and Stalin, Nansen's integrity and stature enabled him to achieve much and brought him the Nobel peace prize in 1922.

A few years before his death in 1930 he was elected rector of St Andrew's University in Scotland. In his inaugural address he fondly reminisced about his Arctic days and quoted Henrik Ibsen's dictum "that man is strongest who stands alone". Huntford brings vividly to life the enigma of a man who stood alone in so many ways as a giant among men but was unsure enough of himself to describe his nature as "a chaos of disharmony".

Nansen was not a leader of men but he

could, when he chose, inspire them. Having made an Arctic journey without equal, he opened the age of modern polar exploration and inspired his successors. He was the stuff of Columbus, Magellan and the Vikings of old. His tragedy was that he was born out of place and out of time.

Nansen emerges, as have so many of the great explorers, as a very complex character. He was much more than the "explorer as hero" of the subtitle. Of Nansen it could surely be more accurately said "the elements so mixed in him that nature might stand up to all the world and say, 'This was a man!'". □

Derek Fordham is at 66 Ashburnham Grove, Greenwich, London SE10 8UJ, UK.

Mental devices

Evolution in Mind

by Henry Plotkin

Allen Lane: 1997. Pp. 276. £25

Massimo Piattelli-Palmarini

The name of the game in cognitive science today is evolutionary psychology. In addition to Henry Plotkin's new book, recent years have seen the publication of many accomplished popularizations of the field, including Steven Pinker's *How the Mind Works*, Terrence Deacon's *The Symbolic Species*, Michael Gazzaniga's *Nature's Mind*, Daniel Dennett's *Darwin's Dangerous Idea* and *Kinds of Minds*, and Richard Dawkins' *Climbing Mount Improbable*. These authors cite one another frequently, being close allies and, in some cases, personal friends. The list is by no means exhaustive, and ignores professional papers in the specialized journals.

Most of these books are well crafted, often crisp, at times peppy, and occasionally infuriating to cognitive scientists of a differ-

ent persuasion. The authors have created a scientific-literary genre, as well as a new subfield of psychology. Academics in other specializations are advised to leaf through at least one of these volumes before going to lunch at the faculty club. Would it not be mildly embarrassing to be found ignorant of evolutionary psychology?

Evolution in Mind is as good a point of entry as any, indeed it has been written precisely for this purpose. Plotkin is a leading exponent of the philosophical school known as 'evolutionary epistemology' (that is, the evolution of knowledge) and has proved in the past, in more scholarly works, to be the kind of experimental psychologist who feels perfectly at ease conversing with the giants of philosophy. His new work is terse and engaging, avoiding not only details and technicalities but also sensationalism and flippancy. It is, in a way, a concise encyclopaedia of psychology — if by psychology one means the natural science of the mind, with a strong emphasis on the word natural.

According to Plotkin, the theory of evolution is the Atlantis of psychology. After the death of William James, a whole Darwinian continent has been lost for psychology, and has to be made to resurface. The evolutionary geneticist Theodosius Dobzhansky used to say that nothing in biology makes sense except in the light of evolution. Plotkin modifies this motto slightly, claiming that "in psychology nothing makes *complete* sense except in the light of evolution" (his emphasis). By judiciously grafting selfish-gene theory, game theory, behavioural ecology and neural networks onto mainstream cognitive science, he provides a fine introductory textbook not only to evolutionary psychology but also to modern psychology in general.

I, for one, siding more with structurally oriented modularity theorists such as Noam Chomsky and Jerry Fodor (who hold that our cognitive abilities are products of innate organs of mind, and whose work the book brilliantly expounds), disagree with Plotkin on only two crucial points. First, because the modularity of mind seems to be not only true, but also arguably (as Plotkin stresses) the most important discovery of modern cognitive science, there is no justification for his alignment with authors who for some reason desire to go 'beyond' modularity, even at the price of smuggling back useless remnants of the defeated theory of 'general intelligence'.

Second, we are still waiting for an example of genuine psychological explanation that descends easily from evolutionary insights but would remain inaccessible to mainstream structural-computational cognitive science. Leda Cosmides and John Tooby, two largely unsung heroes of Plotkin's book, have tried hard to account for our strange and spontaneous conditional

logic by invoking an adaptation-driven 'cheater-detector algorithm'. But their approach, surely novel and stimulating, never did manage to beat more traditional alternatives. The evolutionary psychologists are still in desperate need of patented, specifically Darwinian, mental devices to cash in their vaulting ambitions. □

Massimo Piattelli-Palmarini is in the Department of Cognitive Science, Istituto San Raffaele, Via Olgettina 60, Milano 20132, Italy.

Shades of Green

The Doubly Green Revolution: Food for All in the 21st Century

by Gordon Conway
Penguin: 1997. Pp. 335. £9.99 (pbk)

N. W. Simmonds

This is a serious and substantial book by a thoughtful and well-informed author. An entomologist with 30 years experience, much in the Asian tropics, Gordon Conway is now head of the Institute of Development Studies at the University of Sussex and will soon become president of the Rockefeller Foundation in New York. He therefore occupies

prominent niches in development organizations.

Starting from the historical achievements of the 'green revolution', the book's general theme is how we can feed the human population in the next few decades. The green revolution is well enough described and the author avoids the vulgar error of writing about 'high-yielding varieties' (HYV in the jargon); he recognizes the fundamental genotype-environment interaction but avoids the plant-breeding terminology. He states clearly the essentially local nature of the achievement and the impossibility of duplicating it in any serious sense in other places, times and crops. In future, he rightly says, any major advance must be 'doubly green'. He seems to assume, but does not plainly state, that we must 'buy' enough time for the demographic transition to take over and for food supply to catch up with population growth.

The task is difficult: human population, predominantly in developing countries, is still increasing faster than ever, even though the rate of increase has declined slightly. The 750 million people now chronically underfed will have increased by about another 2,500 million by about 2020, unless agriculture is extraordinarily successful. The cru-

cial requirement will be a vast and sustained rise in yields in diverse crops in diverse environments. The human needs far exceed the good but small achievements of the 'green revolution' with two cereals grown in favoured areas with high inputs of fertilizer and pesticides.

Conway is an enthusiast for concentrating on the needs of the poorest sectors of societies and for farmer 'participation'. The latter has long been recognized as sensible and valuable but is now attaining cult status. 'Farming systems research' was the phrase of the 1970s and 1980s; Conway evidently understood and practised it from an early date but does not call it that any more. Indeed he seems not to have read much of the literature of the 'common goods system', although he praises the achievements. Conway well recognizes the many and heavy side issues that are either already there or can only grow with time. They include global warming and its many uncertainties, pollution, soil erosion, water conservation and use, wood supplies, 'sustainability' (a politicized term), land tenure, communications, food aid and the tragedy of the misused 'common goods'.

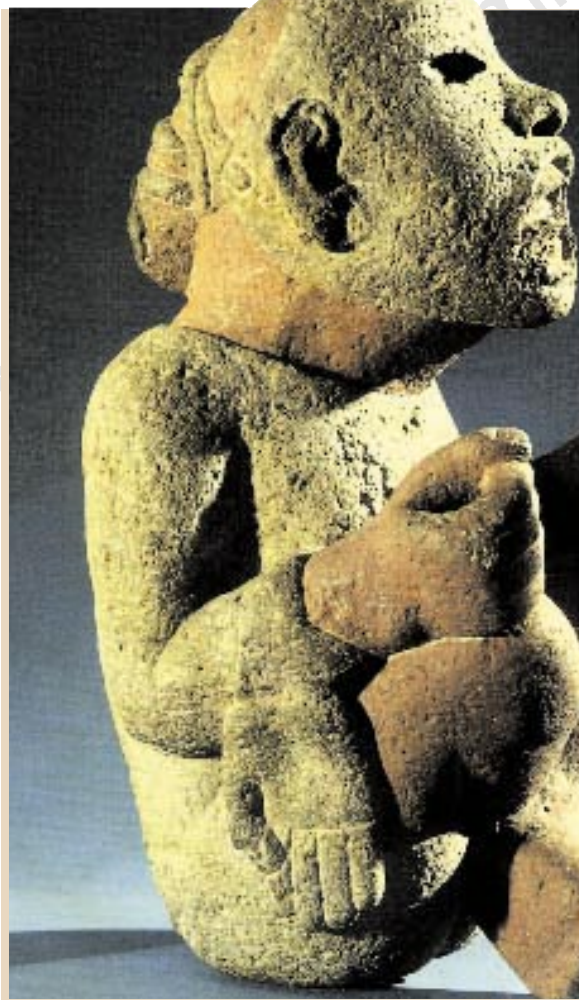
The author is undoubtedly experienced, orders his book well and writes lucidly. The figures are mostly good and clear, but a few are trivial. Although the book is on the whole good, I am unhappy about several fairly weighty matters. First, Conway is an uncritical enthusiast for 'biotechnology', without noticing that its plant breeding achievements so far are nil and that magic is unlikely. He even writes about 'designer plants and animals' as though new varieties were jeans. Second, he ignores the place and potential of cash as a complement to food, and the thought that small farmers might profit from rubber, palm oil or sugar seems as foreign to him as it does to the 'common goods system'.

Third, he recognizes the importance and value of competent agricultural research but does not seem to have read Derek Tribe's books and, apart from participatory enthusiasm, does not reveal how research is to be effectively exploited (even if it gets done, which today seems ever less likely).

Fourth, Conway knows that minerals are essential and that what gets exported from the farm must be replaced, but he does not quite face the fact that fertility comes out of polythene bags and that mucking about with animal composts is, at best, marginal. Here, of course, he hits 'sustainability' (or something) head on because he well knows what some Greens don't yet seem to have learned, that high yields demand high inputs.

If people are ever going to eat more-or-less decently, there will have to be limits to eco-friendliness (which is not to excuse diverse environmental outrages now in progress). □

N. W. Simmonds is at 9 McLaren Road, Edinburgh EH9 2BN, UK.



Getting under their skins

This sculpture shows Xipe Totec, the 'Flayed One', the Aztec god of spring renewal, dressed in the facial and body skin of a sacrificial victim. In Aztec rituals, priests impersonated the god by wearing the

skins of flayed victims. "The bizarre yet apt metaphor of a flayed skin for spring seems quintessentially

Aztec: from death and macabre horror life is reborn." The Almanac calendar dictated the timing of rituals. From *Spirits of the Jaguar: The Natural History and Ancient Civilizations of the Caribbean and Central America* by Paul Reddish (BBC/Parkwest, £18.99, \$32.95), the book that accompanies the BBC television series of the same name.

In retrospect chosen by Owen Gingerich

De revolutionibus orbium celestium ("On the revolutions of the heavenly spheres")

by Nicolaus Copernicus
(1543)

"You may have heard about the thesis of this book and are afraid that all the liberal arts are about to be thrown into confusion," proclaims the introduction to this book. "But these hypotheses need not be true nor even probable.... Perhaps philosophers will seek after truth, but an astronomer will just take what is simplest and neither will find truth unless it has been divinely revealed to him. If you expect to find truth here, beware, lest you leave a greater fool than when you entered."

For most sixteenth-century readers of *De revolutionibus* these were the words of one Nicolaus Copernicus of Torun, canon of the cathedral in the northernmost diocese of Poland, and an astronomer to be reckoned with. Only a few discerned otherwise, and only later was the introduction's anonymous author identified as the Lutheran clergyman Andreas Osiander. While acting as proofreader, Osiander helped the printer to add this protective boilerplate. Misleading as it was in representing the views of Copernicus, it did save the book from condemnation for 73 years. Then, in 1616, Galileo's publicity precipitated action by the Roman Inquisition, which put *De revolutionibus* on the *Index of Prohibited Books* "until corrected". The ten corrections, announced in 1620, were primarily to enhance the hypothetical stance advocated by Osiander.

Judging by the writing in the margins of extant copies, many readers managed to get only about 6% of the way into this formidably technical volume — this being the part that introduced the new heliocentric cosmology, which argued that "Only in this arrangement do we find a sure commensurability between the size of the orbit and its period."

A specialist minority skipped over the opening section and delved deeply into the remaining 94%, filled with the geometry to



solve solar, lunar and planetary motion. There they found that Copernicus, driven by the aesthetic nature of uniform speed on perfect circles, had used a series of epicycles to accomplish his goal. It is hard for us to imagine anyone getting steamed up by this sort of thing, and we can nod sympathetically with Arthur Koestler's declaration that this was the book nobody read. In fact, Koestler was quite wrong. There were surprisingly many readers, though rather few gave the technical passages intense scrutiny.

Despite the impression given by innumerable secondary accounts, Copernicus had no observational proof of his heliocentric arrangement. It was an achievement in the mind's eye, a theory "pleasing to the mind". But it accorded in a reasoned way with many natural phenomena, which in the Sun-centred plan were "linked together as by golden chain", in the words of his young disciple, Georg Joachim Rheticus. As Copernicus pointed out, his scheme explained quite naturally why the retrograde motion of nearby Mars was larger than that of more distant Jupiter, and why the still more distant stars themselves failed to show this annual backward wobble. "So vast, without any question, is the

divine handiwork of the most excellent Creator." This was one of the passages censored by the Roman Inquisition. A pious statement, they thought, it exuded too great a whiff of reality, as if that is how God actually made the cosmos.

But did Copernicus actually believe in the physical reality of his system? Most probably yes. But elements of the later, technical parts of *De revolutionibus* leave a lingering doubt. For example, on the planetary latitudes Copernicus uses mechanisms entirely different from those he has earlier presented for the longitude. But surely a physically real system would demand the same mechanism for both latitudes and longitudes, and this is what the thorough-going realist Kepler eventually demanded, and which forced Copernicus onto the path that led to the elliptical orbit and the law of areas. Nevertheless, the vision of Copernicus formed the blueprint on which newtonian mechanics and the modern world would eventually be built.

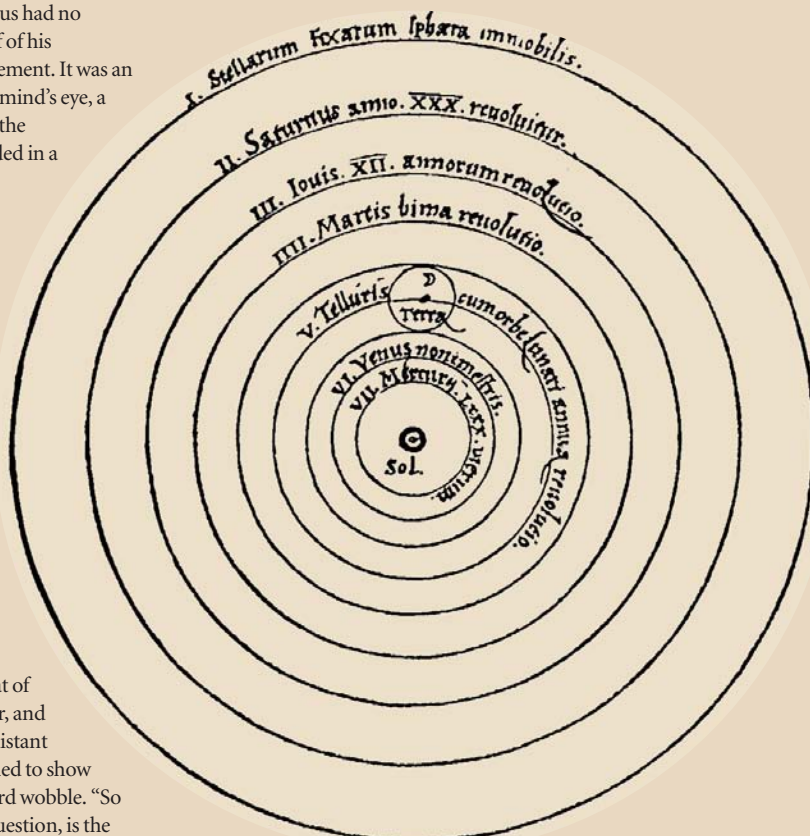
Early on, *De revolutionibus* had a steady, if small, stream of readers.

The work then languished, made more or less obsolete by Kepler's *Epitome of Copernican Astronomy*.

In 1939 an English edition, not altogether reliable, finally appeared. In 1976, Alistair Duncan's translation came out, followed by Edward Rosen's version in 1978. Rosen's is the more literate, especially in the early cosmological sections, whereas Duncan seems more reliable in the lengthier, technical parts. As an aid to understanding those technical parts, *The Mathematical Astronomy of Copernicus's De Revolutionibus* by Noel Swerdlow and Otto Neugebauer is unsurpassed.

Today the 1543 *editio princeps* of *De revolutionibus orbium celestium* has become an icon of scientific progress, the first great work of the astronomical revolution. As such, copies now fetch an astronomical price, up to US\$300,000. It is wonderful that achievements of the intellect, theories "pleasing to the mind", are now so highly regarded.

Owen Gingerich is at the Harvard-Smithsonian Center for Astrophysics, 60 Garden St, Cambridge, Massachusetts 02138, USA.



Timing of abrupt climate change at the end of the Younger Dryas interval from thermally fractionated gases in polar ice

Jeffrey P. Severinghaus*, Todd Sowers†, Edward J. Brook‡, Richard B. Alley† & Michael L. Bender*§

* Graduate School of Oceanography, University of Rhode Island, Narragansett, Rhode Island 02882, USA

† Geosciences Department, Pennsylvania State University, University Park, Pennsylvania 16802, USA

‡ Departments of Geology and Environmental Science, Washington State University, Vancouver, Washington 98686, USA

Rapid temperature change fractionates gas isotopes in unconsolidated snow, producing a signal that is preserved in trapped air bubbles as the snow forms ice. The fractionation of nitrogen and argon isotopes at the end of the Younger Dryas cold interval, recorded in Greenland ice, demonstrates that warming at this time was abrupt. This warming coincides with the onset of a prominent rise in atmospheric methane concentration, indicating that the climate change was synchronous (within a few decades) over a region of at least hemispheric extent, and providing constraints on previously proposed mechanisms of climate change at this time. The depth of the nitrogen-isotope signal relative to the depth of the climate change recorded in the ice matrix indicates that, during the Younger Dryas, the summit of Greenland was $15 \pm 3^\circ\text{C}$ colder than today.

Abrupt changes in the Earth's climate system have generated much interest recently. Ice-core studies that infer palaeotemperature from the ratio of oxygen isotopes in the ice^{1,2} and accumulation rate³ have suggested that Greenland temperatures rose in less than a decade at the climate transition marking the end of the Younger Dryas cold interval and the beginning of the warmer Holocene epoch at 11.6 kyr before present, BP (here 'present' indicates AD 1950). Abrupt changes of similar age have appeared in other climate records over much of the globe. Whether these widespread changes were simultaneous, asynchronous, or unrelated has significant implications for our understanding of them, but dating uncertainties a century or more at present obscure these relationships.

One critical record of Younger Dryas climate change, also derived from ice cores, is the atmospheric methane concentration in trapped air. Analysis of the air trapped in bubbles in ice from Greenland^{4–6} and from the Antarctic⁷ has shown that atmospheric methane concentrations also increased near the end of the Younger Dryas. It is thought that this increase was in response to increased precipitation^{4,5} and/or temperature⁵ in the wetland methane source regions, which at that time were widely distributed over the tropics and possibly the middle- to high-latitude Northern Hemisphere^{4,5,7}. Because the atmosphere is well mixed, its methane concentration serves as an integrator of methane sources (and thus climate information) over a broad spatial scale. As few methane sources existed in the North Atlantic basin during this period⁵, a climate change that was restricted to this area alone could not have produced the observed atmospheric methane changes. Knowledge of the precise relative timing of the changes in temperature and methane should thus show whether the postulated abrupt Greenland temperature change was synchronous with climate change over a large portion of the Earth's terrestrial area. This information may in turn help constrain causal mechanisms of the abrupt changes.

Two problems currently prevent us from relating changes in temperature and methane in the ice core. First, the relationship between the $^{18}\text{O}/^{16}\text{O}$ ratio of ice ($\delta^{18}\text{O}_{\text{ice}}$) and the palaeotemperature

has been shown to vary with age^{8,9}, and we do not know the exact relationship at times of rapid climate change during the ice ages^{10,11}. Recent borehole temperature studies⁸ show that the $\delta^{18}\text{O}_{\text{ice}}$ 'palaeothermometer' underestimates the glacial-to-Holocene warming by about a factor of 2 when calibrated from the modern spatial $\delta^{18}\text{O}$ –temperature relationship, motivating the search for an independent palaeothermometer. Second, the age of the air trapped in bubbles is less than the age of the enclosing ice because air is occluded at some depth below the surface of the ice sheet in the bubble close-off region^{12–14}. This gas-age–ice-age difference is not known *a priori* for times past, making uncertain the relative timing of temperature and atmospheric gas changes^{4,14}.

Here we address these problems with a new technique for identifying rapid past temperature changes based on the principle of thermal diffusion, which fractionates gas mixtures in a temperature gradient according to their mass. The air in the porous unconsolidated layer of snow (firn) on top of polar ice sheets is fractionated by transient temperature gradients arising from abrupt climate change. This fractionated air is preserved in bubbles in glacial ice as a gas-isotope anomaly and thus serves as a gas-phase stratigraphic marker of the temperature–change event in an ice core. This approach circumvents the problem of the gas-age–ice-age difference by comparing gases directly with gases, enabling a precise determination of the relative time of local temperature changes and atmospheric gas changes. We apply the method to the GISP2 ice core from the summit of Greenland using the stable isotopes of N_2 and Ar in trapped air, and find that the increase in atmospheric methane concentrations at the end of the Younger Dryas interval began 0–30 years after an abrupt temperature increase. Where annual layers are countable¹⁵, as in GISP2, the position of the gas-isotope anomaly relative to the record of the temperature event in the ice yields a precise estimate of the gas-age–ice-age difference, which allows a tie between atmospheric gas chronologies and the layer-counting timescale at times of abrupt change. We also use this age difference in conjunction with glaciological models^{16,17} (which predict the depth and age of pore close-off in the firn) to obtain an estimate of absolute temperature that serves as a check

§ Present address: Department of Geosciences, Princeton University, Princeton, New Jersey 08544, USA.

on the $\delta^{18}\text{O}_{\text{ice}}$ palaeothermometer. Finally, the relative changes in $^{15}\text{N}/^{14}\text{N}$ and $^{40}\text{Ar}/^{36}\text{Ar}$ confirm that the abrupt shift in $\delta^{18}\text{O}_{\text{ice}}$ at the Younger Dryas/Holocene transition was indeed a thermal event.

Thermal fractionation of air in the firn

A gas mixture subjected to a temperature gradient will tend to unmix by a process known as thermal diffusion^{18,19}. The effect was predicted with the solution of the Boltzmann equation and the development of the kinetic theory of gases in 1911–17, which showed that diffusive molecular transport could be driven by a temperature gradient (thermal diffusion) in addition to a concentration gradient (concentration diffusion)¹⁸. Subsequently confirmed in the laboratory²⁰, the effect was studied extensively during the 1940s as a means to separate uranium isotopes. In a constant temperature gradient, a steady state is reached in which thermal diffusion in one direction is balanced by diffusion along a concentration gradient in the other direction. In this situation the difference in the isotope ratios R and R_0 at temperatures T and T_0 is given (in units of per mil) by theory as¹⁹

$$\delta = [R/R_0 - 1]10^3 = \left(\left[T_0/T \right]^\alpha - 1 \right) 10^3 \quad (1)$$

where δ is the fractional deviation of the isotope ratio R from the ratio R_0 , T and T_0 are temperature in K, and α is the thermal diffusion factor¹⁸, a parameter characteristic of a pair of gas species. With a few exceptions, the heavier gas becomes enriched in the colder region¹⁸. For example, α for the isotopic pair $^{15}\text{N}/^{14}\text{N}$ – $^{14}\text{N}_2$ is about 0.0065 at 298 K (ref. 18), so a temperature gradient between 298 and 308 K will cause the cold end to have $\delta^{15}\text{N} = +0.2\text{‰}$ relative to the warm end. For comparison, our measurement precision for $\delta^{15}\text{N}$ in ice cores is $\pm 0.02\text{‰}$ corresponding to $\pm 1^\circ\text{C}$. In our work, the gas at one end of the temperature gradient is always the free atmosphere at the top of the firn. The atmosphere is used as the reference for both $^{15}\text{N}/^{14}\text{N}$ and $^{40}\text{Ar}/^{36}\text{Ar}$ measurements, as these ratios are constant in air on the 10^4 -year timescale relevant here^{14,21}.

A second process known as gravitational settling^{22,23} also affects the gas composition in firn^{14,24,25}. Separation of gases in an isothermal, porous column occurs according to the mass difference Δm alone as given by the barometric equation²⁴

$$\delta = [R/R_0 - 1]10^3 = \left(\exp \left[\Delta mgz / (R^* T) \right] - 1 \right) 10^3 \quad (2)$$

where R is the isotopic ratio at depth z in the column, R_0 is the ratio in air, g is the gravitational acceleration, R^* is the gas constant and δ is in units of per mil. For example, in a diffusive column of air 80 m thick at 236 K, the air at the bottom will have $\delta^{15}\text{N} = +0.4\text{‰}$ relative to the air at the top.

Observations of thermal diffusion in nature have thus far been limited to gas-filled porous media such as sand dunes²⁶ and firn²⁷, where advective mixing is inhibited by the small pore size such that transport is mainly by diffusion. These observations confirm that transient seasonal^{26,27} and geothermal²⁶ temperature gradients fractionate air approximately as predicted.

After an abrupt climate warming, transient temperature gradients lasting several hundred years should arise in the upper part of the firn owing to the thermal inertia of the underlying firn and ice. The heavy gas species such as $^{15}\text{N}/^{14}\text{N}$ and ^{40}Ar should therefore be preferentially driven downwards, towards the cold deeper firn, relative to the lighter species ($^{14}\text{N}_2$ and ^{36}Ar). An important point is that gases diffuse about 10 times faster than heat in firn²⁸, so the thermally fractionated gas will penetrate to the bottom of the firn long before the temperature equilibrates, and the gas composition will closely approach a steady state with respect to the firn temperature such that equation (1) should be valid. Figure 1 shows model results of the expected firn-air composition in the 50 years following a 5°C 'step function' increase in surface temperature (see Methods). We note that the isotopic signal propagates to the bottom

of the firn during the first decade, superimposed on the linear downward increase in $\delta^{15}\text{N}$ due to gravitational settling. By 50 years after the initial warming, $\delta^{15}\text{N}$ at the bottom of the firn rises to the steady-state value of thermal fractionation (the signal is not attenuated at this time). As the bubbles close off they trap a sample of the basal firn air, thus recording the thermal event as a heavy-gas anomaly in the bubble air composition. Changes in atmospheric methane concentrations should also propagate downwards at the same speed, with a slight correction for the fact that CH_4 diffuses $\sim 7\%$ faster than $^{15}\text{N}/^{14}\text{N}$. These CH_4 variations will be trapped in the bubbles along with the thermal signal, allowing a precise estimate to be obtained of the temporal relationship between the abrupt change

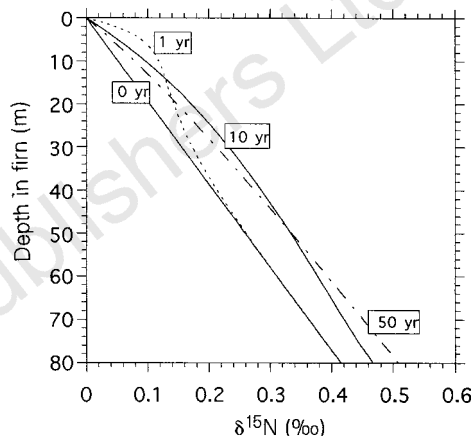


Figure 1 Results of a heat and molecular diffusion model of gas in polar firn. This model was used to simulate the effect on firn-air isotopic composition of a 5°C 'step function' increase in surface temperature (constant temperature before and after the step). $\delta^{15}\text{N} = [(^{15}\text{N}/^{14}\text{N})_{\text{sample}} / (^{15}\text{N}/^{14}\text{N})_{\text{atmosphere}}] - 1 \times 10^3$, where $\delta^{15}\text{N}$ is in units of per mil. The time in years after the step is shown by the numbers in boxes. See Methods section for model architecture.

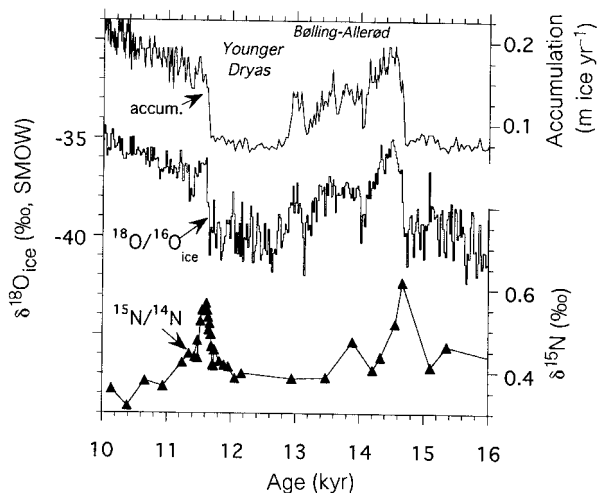


Figure 2 GISP2 records of accumulation⁹, $\delta^{18}\text{O}_{\text{ice}}$ (ref. 2), and $\delta^{15}\text{N}$ of trapped air (ref. 47 and this work) covering the last deglaciation. Age values for $\delta^{15}\text{N}$ were obtained from a model estimate¹⁷ of the gas-age-ice-age difference and may be in error by ~ 100 yr. Means of replicate $\delta^{15}\text{N}$ measurements are shown. Note the $\delta^{15}\text{N}$ anomaly of about $+0.15\text{‰}$ at 11.6 kyr BP, followed by a decline after several hundred years. An even larger anomaly appears to be present at the beginning of the Bølling-Allerød and is the subject of current work. The abrupt cooling at the beginning of the Younger Dryas is expected to produce a negative anomaly in $\delta^{15}\text{N}$, the apparent absence of which may reflect low sampling resolution in that interval.

in surface temperature and the CH_4 increase. Further modelling shows that after several hundred years the thermal diffusion anomaly decays as the firn becomes isothermal. In the case of a warming, the $\delta^{15}\text{N}$ anomaly observed in ice cores should thus take the form of a sudden increase followed by a gradual decrease. Sudden cooling is expected to produce the opposite effect.

Observations of nitrogen and methane in trapped air

As can be seen Fig. 2, $\delta^{15}\text{N}$ in the GISP2 ice core during periods of cold, stable climate is about $+0.4\text{‰}$, the expected value for gravitational settling with a firn thickness of $\sim 80\text{ m}$. The end of the Younger Dryas interval at 11.6 kyr BP is identified by the abrupt

increase in $\delta^{18}\text{O}_{\text{ice}}$. At this point the $\delta^{15}\text{N}$ record shows a 'spike' to $+0.56\text{‰}$, followed by a decay over several hundred years. A similar anomaly is seen in the $^{18}\text{O}/^{16}\text{O}$ ratio of O_2 (not shown), with twice the amplitude of the $\delta^{15}\text{N}$ anomaly, as expected from reported values of the thermal diffusion factor¹⁸. Figure 3a shows an expanded view of the $\delta^{15}\text{N}$ anomaly plotted versus depth in the ice. We interpret the sudden rise in $\delta^{15}\text{N}$ at $\sim 1,700\text{ m}$ depth as the gas-phase marker of the climate event. By minimizing the mismatch between model $\delta^{15}\text{N}$ (Fig. 4) and the data, we assign the event age³ of 11.64 kyr BP to the gas at $1,700.3\text{ m}$, with an estimated depth uncertainty corresponding to $\pm 20\text{ yr}$ (based on the sample resolution).

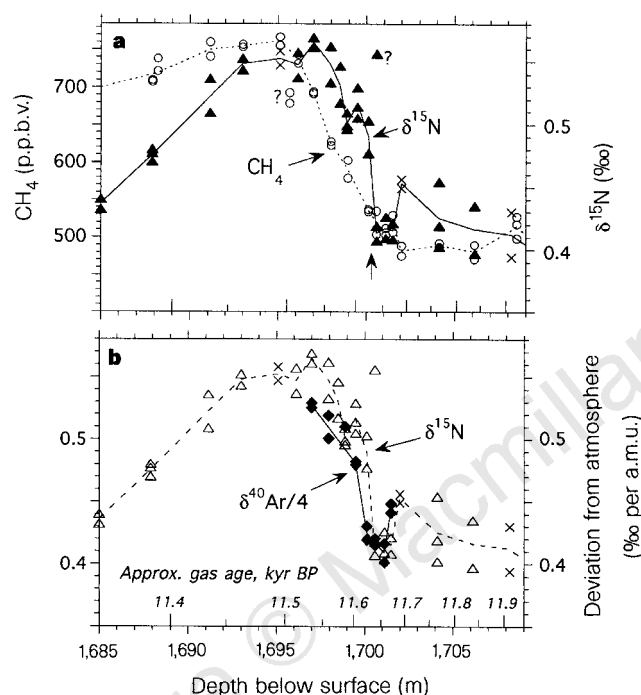


Figure 3 Gas data from GISP2 covering the end of the Younger Dryas, plotted versus depth. **a**, Expanded view of the $\delta^{15}\text{N}$ anomaly in Fig. 2 showing individual replicates, with methane data superimposed for comparison. Lines pass through the means of replicates, except for a pair of methane points at $1,695.7\text{ m}$ which we consider to be anomalous. We reject the $\delta^{15}\text{N}$ outlier at $1,700.6\text{ m}$ and draw the line through the mean of the remaining two replicates. Cross symbols (x) denote $\delta^{15}\text{N}$ measurements made in an earlier study⁴⁷. Note the sharp increase in $\delta^{15}\text{N}$ at $\sim 1,700\text{ m}$ depth from the baseline value of $+0.4\text{‰}$ to $+0.5\text{‰}$. Samples spanning this change are 18 yr apart, placing an upper limit on the amount of smoothing and hence the duration of the air enclosure process. We interpret the inflection point in $\delta^{15}\text{N}$ at $1,700.3\text{ m}$ (arrow) as a marker of the thermal event in the gas phase. The $\delta^{15}\text{N}$ outlier at $1,700.6\text{ m}$ makes the location of the inflection point ambiguous, so we accept a location error corresponding to one sample spacing. For reference, approximate gas age is plotted along the horizontal axis, adjusted to be consistent with our conclusion that the $\delta^{15}\text{N}$ anomaly began at 11.64 kyr BP . **b**, $\delta^{40}\text{Ar}/4$ in ice adjacent to the $\delta^{15}\text{N}$ samples, superimposed on the $\delta^{15}\text{N}$ data for comparison. $\delta^{40}\text{Ar} = [(^{40}\text{Ar}/^{36}\text{Ar})_{\text{sample}} / (^{40}\text{Ar}/^{36}\text{Ar})_{\text{atmosphere}}] - 1110^3$, where $\delta^{40}\text{Ar}$ is in units of per mil. $\delta^{40}\text{Ar}$ is divided by four for comparison with $\delta^{15}\text{N}$ due to the four mass unit difference between ^{40}Ar and ^{36}Ar ; gravitational fractionation is the same for all isotopes on a per mass unit basis. Lines connect means of replicate points. Note the rough agreement of Ar and N_2 at depths $>1,700\text{ m}$, implying that gravitation is the only process at work. In contrast, at depths $<1,700\text{ m}$ $\delta^{40}\text{Ar}/4$ data fall below $\delta^{15}\text{N}$ data, with the exception of the single (unreplicated) Ar point at $1,698.8\text{ m}$. Integrity of this sample may have been compromised because it came from the outer 2 cm of the core. The lower $\delta^{40}\text{Ar}/4$ requires a thermal fractionation component in the points at $1,697, 1,698, 1,699.5$ and $1,700\text{ m}$. Note unexplained discrepancies of $\sim 0.03\text{‰}$ in samples from $1,700$ and $1,701.5\text{ depth}$ (a.m.u., atomic mass unit).

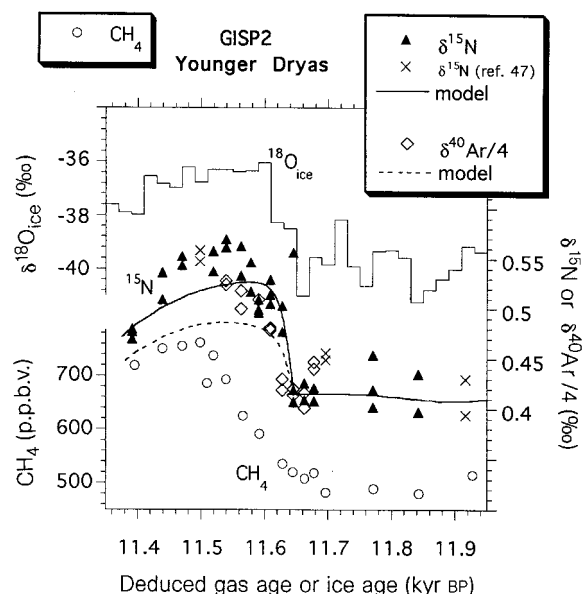


Figure 4 Isotopic and CH_4 data plotted versus age derived by this work. Bidecadal mean $\delta^{18}\text{O}_{\text{ice}}$ (ref. 2) is shown for comparison (top). Model $\delta^{15}\text{N}$ (solid line) is the result of a simulation forced by a 5°C linear increase in surface temperature over the 10-yr period from 11.64 to 11.63 kyr BP , with constant temperature before and after these times. Only the relative changes in the model have significance; model absolute values are arbitrary and are forced to go through the mean ($+0.415\text{‰}$) of the six observed $\delta^{15}\text{N}$ values at 11.64 – 11.68 kyr BP (the outlier at 11.64 kyr BP is excluded). The gravitational effect of column deepening is added to the model curve using model¹⁷ predicted firn thickness changes (this effect reaches a maximum of $+0.034\text{‰}$ per mass unit at 11.5 kyr BP). The age scale was derived by assuming that the gas age at $1,700.3\text{ m}$ is 11.64 kyr BP , and calculating ages relative to this point from gas-age-ice-age differences predicted by a dynamic densification model¹⁷.

Methane concentration analyses⁵ of air trapped in bubbles are superimposed for comparison in Fig. 3a. Methane concentration first rises significantly beyond the envelope of Younger Dryas variability (± 30 parts per billion by volume, p.p.b.v.) at 1,699 m depth, which is ~ 30 years after the corresponding increase in $\delta^{15}\text{N}$ at 1,700.1 m depth. Because there are no CH_4 samples between these depths, we infer that atmospheric methane began to rise 0–30 years after the warming event. Allowance for the slightly faster diffusion of CH_4 than $^{15}\text{N}/^{14}\text{N}$ makes the zero-year lag slightly less likely. Importantly, methane rises more gradually than $\delta^{15}\text{N}$, taking >150 years to reach maximum values. In contrast, modelling of the rapid $\delta^{15}\text{N}$ change between 1,700.6 and 1,700.1 m suggests a step-like temperature change that occurred in less than a decade (Fig. 4).

Separation of thermal and gravity effects

It is likely that the $\delta^{15}\text{N}$ anomaly at 11.6–11.4 kyr BP is partly due to a sudden deepening of the diffusive firn column and the resulting increase in gravitational settling. A transient deepening of the firn is expected owing to the doubled accumulation rate³ at this time. A dynamic densification model¹⁷ predicts 6 m of deepening, which would increase $\delta^{15}\text{N}$ by 0.03‰. Wind speed also may have dropped at this time^{29,30}, which could have decreased the thickness of a near-surface convective zone³¹ in which gravitational separation is prevented by wind-pumping³², thereby deepening the diffusive column. To address the issue of how much of the isotopic signal is due to temperature change, we exploit the fact that Ar is half as sensitive to thermal fractionation as N_2 (ref. 18). This allows us to separate thermal and gravitational effects by measurement of $^{40}\text{Ar}/^{36}\text{Ar}$ and Ar isotopes in the same ice. If the $\delta^{15}\text{N}$ anomaly is due entirely to increased gravitational settling, then Ar isotopes should show an anomaly with the same amplitude as N_2 isotopes, when scaled by mass difference (that is, $\delta^{40}\text{Ar}/4$ should equal $\delta^{15}\text{N}$). On the other hand, if the anomaly is due entirely to thermal fractionation then $\delta^{40}\text{Ar}/4$ should change only half as much as $\delta^{15}\text{N}$.

Figure 3b shows $\delta^{40}\text{Ar}/4$ from ice samples at the same depths as the $\delta^{15}\text{N}$ samples. The anomaly in Ar appears to be less than that of N_2 , and therefore we conclude that the anomaly marks a thermal fractionation event. Because the Ar amplitude is about 3/4 rather than 1/2 of the N_2 amplitude, the anomaly must also be partly due to gravitation via diffusive column deepening. A preliminary quantification of the thermal and gravity components gives 5–10 °C of abrupt warming, possibly followed by a gradual column deepening over a century. We consider these estimates highly tentative.

Analytical uncertainties are large, and the thermal diffusion factors for N_2 and Ar at -40 °C remain to be measured. The possibility remains that part of the abrupt $\delta^{18}\text{O}_{\text{ice}}$ increase at 11.6 kyr BP was due to a geographical shift^{11,33–35} in the location of the moisture source rather than a warming at the drill site (summit).

Climatic implications of methane trends

Using the $\delta^{15}\text{N}$ anomaly as a stratigraphic marker of the warming event, we conclude that the methane increase began 0–30 years after the warming at Summit. Subsequently, however, methane increased more slowly than the temperature such that the bulk of the methane increase came after the warming (Fig. 4). Therefore a causal role for methane in the warming, already doubtful owing to the small greenhouse effect of methane⁶, is clearly ruled out by these phase data. Instead, methane seems to be responding to the climate change. The dominant source of atmospheric methane in the pre-industrial period was wetlands³⁶. Because changes in sink strength are thought to have been small³⁷, atmospheric concentration variations are thought to primarily reflect variations in wetland temperature⁵ and/or wetland area and hence terrestrial precipitation minus evaporation^{4,36}. Because the residence time of methane in the atmosphere is about a decade³⁸, changes in the methane source could have preceded the observed atmospheric change by no more than a decade. As noted previously (but with less resolution)^{4,5}, the near synchronicity of the methane and temperature changes and the widespread distribution of methane sources imply that this climate event was at least hemispheric in extent. Furthermore, the speed of the apparent link between methane source regions and Greenland is evidence supporting an atmospheric transmission of the climate signal, as transmission through the ocean would have been too slow.

Our results may have implications for a proposed link³⁹ between changes in the tropical hydrological cycle and North Atlantic deep water (NADW) formation. An increase in the net evaporation rate over the tropical Atlantic might increase the salinity of poleward-bound subtropical surface water and hence increase NADW formation³⁹, as well as increasing precipitation over land⁴. An increase in NADW formation is widely thought to have caused the abrupt Greenland warming⁴⁰. If an increase in tropical precipitation induced the atmospheric methane rise, methane might be expected to lead the Greenland warming by the amount of time required for saline water to make its way from the tropics to the North Atlantic⁴ (several decades). Our results do not provide

Table 1 Palaeothermometry based on observed gas-age–ice-age difference

	Input parameters				Greenland summit temperature (°C)			
	Δage (yr)	Accumulation ($\text{m H}_2\text{O yr}^{-1}$)	Density		This work	Measured	$\delta^{18}\text{O}$ thermometer	
			Lock-in (kg m^{-3})	Initial (kg m^{-3})			Borehole-based (ref. 8)	Spatial gradient-based
Present (GLSP2 site, ref. 17)	195	0.227	814	350	–30	–31		
Younger Dryas (GLSP2 site)	809*	0.073†	820‡	360§	–46		–47	–39
1 σ error	± 20	± 0.01	± 15	± 10	± 3			
Sensitivity test for lock-in density	809	0.073	810	360	–47.3			
	809	0.073	835	360	–43.7			

The last two columns at the right show, for comparison, the temperatures estimated using the traditional oxygen isotope palaeothermometer for two different calibrations ($d\delta/dT$) based on borehole thermometry (0.33) and the modern spatial covariation of $\delta^{18}\text{O}$ with temperature (0.65). The latter neglects changes in seawater $\delta^{18}\text{O}$ due to ice volume change. The input parameters chosen imply that Younger Dryas lock-in depth was 92 m (ref. 16). For comparison, the height of the firn diffusive column inferred from the Younger Dryas mean $\delta^{15}\text{N}$ of $+0.415\text{‰}$ (Fig. 3) and equation (2) is 80 m. The discrepancy could be due to the presence of a $\sim 12\text{-m}$ -thick near-surface zone that was well flushed by wind-pumping (refs 31, 32).

* This work.

† Ref. 9. This is the mean value for the last 400 years of the Younger Dryas, for the scenario of 200 km of net ice margin retreat during deglaciation. Error is subjective and is based on model scenarios (K. Cuffey, personal communication).

‡ Calculated from ref. 17; also equal to South Pole present value (ref. 42). Error is subjective and is based on the range of published estimates.

§ Value used in Fig. 2 in ref. 16. Error is subjective. The result is nearly insensitive to this parameter.

|| Ref. 16 includes contribution from scatter in the Herron-Langway model (Fig. 1 in ref. 16). The model predicts modern temperature with 1 σ error of ± 2 °C (see text). To include this error in our result, we assign an error to the prefactor (575) of the rate constant k_1 in ref. 16. We then adjust this error until it produces a ± 2 °C error in the result, with all other errors set to zero. The resulting prefactor error is ± 60 , which is then propagated in the Monte Carlo method along with the other four errors.

support for such a link inasmuch as methane does not appear to lead temperature at this one warming event.

Palaeothermometry from stratigraphic marker position

We can also use the $\delta^{15}\text{N}$ anomaly at the beginning of the Younger Dryas termination as a palaeothermometer. Colder temperatures slow the sintering of firn into ice because it is a temperature-activated metamorphic process¹⁶. The age of the ice at which air is trapped depends on firn temperature and ice accumulation rate¹⁶. In GISP2 there are 809 ± 20 annual layers^{3,15} between the beginning of the $\delta^{15}\text{N}$ increase (1,700.3 m depth) and the shift in $\delta^{18}\text{O}_{\text{ice}}$ that marks the warming in the solid phase. An age difference of 809 years and an accumulation rate of $0.073 \text{ m water yr}^{-1}$ (refs 3, 9) imply a temperature of $-46 \pm 3^\circ\text{C}$ (Table 1), which is $15 \pm 3^\circ\text{C}$ colder than at present (this calculation assumes steady-state before the termination). To check the method we use it to estimate modern temperature at Summit; we obtain -30°C which is close to the measured value of -31°C .

This estimated temperature of $-46 \pm 3^\circ\text{C}$ agrees well with the borehole temperature-based estimate of -47°C for this time⁸, but is significantly colder than the value inferred from the spatially calibrated $\delta^{18}\text{O}_{\text{ice}}$ palaeothermometer (ref. 35; Table 1). (We point out that mean tropospheric temperature might not have been 15°C colder than at present if the near-surface atmospheric inversion layers common on polar ice sheets were stronger at Summit during the Younger Dryas interval⁹.)

This 15°C temperature change is not inconsistent with the figure of $5\text{--}10^\circ\text{C}$ noted earlier for the amount of abrupt warming, because of the very different time spans involved. The entire transition from Younger Dryas to typical Holocene $\delta^{18}\text{O}_{\text{ice}}$ values took $\sim 1,500$ years (Fig. 2), over which time the 15°C change probably occurred, whereas the $5\text{--}10^\circ\text{C}$ figure applies to a 'step' change that occurred over several decades or less. In this context we note that the thermal diffusion tool is not sensitive to gradual temperature change ($<0.01^\circ\text{C yr}^{-1}$), because the firn becomes isothermal on a timescale of several hundred years. □

Methods

Analytical. Samples of ice were analysed respectively for N_2 (10-g samples) and Ar (35-g samples) isotopic composition with a melt-refreeze technique that releases the air trapped in bubbles¹⁴. Ice was generally taken from the central portion of the core to minimize possible adverse effects of gas loss due to temperature fluctuations during core retrieval and handling. N_2 isotopes were measured as in ref. 14. Air for Ar analysis was extracted in two melt-refreeze cycles to increase yield, and the air was exposed to a Zr/Al SAES getter at 900°C for 10 min removing all the N_2 , O_2 and other reactive gases. Pure N_2 equal to $10\times$ the residual Ar was then added to increase sample volume to improve mass spectrometry, and the ratio of mass 40 to mass 36 of the mixture was measured on a multi-collector Finnigan MAT 252 mass spectrometer. Corrections of the order of 0.01‰ were made for the sensitivity of the relative ionization efficiencies of ^{40}Ar and ^{36}Ar in the mass spectrometer source to variations in the Ar/ N_2 ratio and to pressure imbalance between sample and reference inlets. The standard deviation from the mean of replicate ice samples cut from the same depth was $\pm 0.02\text{‰}$ for $\delta^{15}\text{N}$, and $\pm 0.03\text{‰}$ for $\delta^{40}\text{Ar}$ (or $\pm 0.008\text{‰}$ on a per mass unit difference basis). Analytical precision for analyses of firn air is about $\pm 0.01\text{‰}$ for both $\delta^{15}\text{N}$ and $\delta^{40}\text{Ar}$, suggesting that the larger variability in the ice-core results was due either to artefacts introduced during the sample extraction procedure or real sample-scale heterogeneities in the ice. Simulated extractions were performed by introducing a sample of dry Rhode Island air to the extraction vessel over thoroughly degassed GISP2 ice, then following the normal extraction procedure. Results of these experiments revealed no systematic bias resulting from the extraction procedure. Methane procedures and results are described elsewhere⁵.

Numerical simulation. A time-dependent gas diffusion model forced by an independent heat advection–diffusion model was used to explore the magnitude and shape of the expected firn-air and ice-core isotopic anomaly resulting from sudden climate change. Advection of molecules was neglected.

Three processes were treated in the gas diffusion model: concentration diffusion (Fick's Law), thermal diffusion, and gravitational settling. Gas fluxes were calculated from the difference between the concentration gradient and the thermal/gravitational equilibrium gradient, in effect relaxing the concentrations back to the equilibrium profile at the rate given by the diffusivity (following ref. 26). Justification for this approach is based on the fact that thermal diffusion proceeds at the speed of concentration diffusion²⁰ and is further discussed in ref. 41. Model gas diffusivities decreased monotonically downward in the firn, and were prescribed based on an assumed density profile and an empirical density–diffusivity relation derived from CO_2 measurements in South Pole firn air⁴². The firn–ice transition was held constant at 80 m and treated as an impermeable boundary condition. The heat model was adapted from ref. 43. Thermal conductivity and heat capacity were scaled to density based on empirical data. The insulating bottom boundary of the heat model was extended to 800 m depth to account for the thermal inertia of the ice. The initial condition of the model was isothermal at 227 K , and the model was forced at the upper boundary condition with a 5°C step function.

Palaeotemperature inversion. Herron and Langway¹⁶ showed that firn densification is a function of temperature and snow accumulation rate, and constructed an empirical model of the densification process based on measured density–depth profiles from 17 locations in Greenland and Antarctica with firn temperatures at 10 m depth ranging from -15 to -57°C and accumulation rates ranging from 0.5 to $0.02 \text{ m water yr}^{-1}$. Using their regression coefficients, we are able to predict modern temperature (T) at these 17 sites from the given accumulation (A) and density–depth ($d\rho/dz$) data to within $\pm 2^\circ\text{C}$ (1σ) of the measured temperatures, suggesting the potential usefulness of the model as a palaeothermometer.

The layer-counting age of our inferred gas-phase temperature change marker at 1,700.3 m is $12,449 \text{ yr}$ (ref. 15), compared to the layer-counting age of $11,640 \text{ yr}$ of the climate event as recorded in the ice by $\delta^{18}\text{O}_{\text{ice}}$ (ref. 2), electrical conductivity⁴⁴ and accumulation³, giving an inferred gas-age–ice-age difference (Δage) of 809 yr . We adopt a subjective uncertainty of $\pm 20 \text{ yr}$ based on the $\delta^{15}\text{N}$ sample spacing of $\sim 18 \text{ yr}$ (Fig. 3a) and the ambiguity introduced into our interpretation of the depth of the $\delta^{15}\text{N}$ inflection point by a single outlier at 1,700.58 m (Fig. 3a). Relative uncertainty in the layer counting in this interval is of the order of 1% or $\sim 8 \text{ yr}$ (ref. 15).

To relate this value of Δage to the Herron–Langway model, we assume that gases are cut off from mixing with the atmosphere at a specified density known as the lock-in density^{13,42} (ρ_L). This assumption is predicated on observations of firn air that show that at most sites air is cut off from mixing with the atmosphere by impermeable high-density winter layers before the bubbles are all closed^{13,25,42,45,46}. We also must specify an initial density at the surface (ρ_0), which has little effect on the result. Accumulation is from Cuffey and Clow's⁹ corrections for thinning to measurements of layer thickness using a coupled ice and heat flow model. We solve equations (11) and (9) in Herron and Langway¹⁶ to obtain T (in K) as an iteratively solved function of Δage , A , ρ_L and ρ_0 :

$$T = \frac{21400}{R^*} \left/ \ln \left\{ \frac{575A^{0.5}}{\ln \left[\frac{\rho_i - 550}{\rho_i - \rho_L} \right]} \left(\left[\Delta\text{age} + \tau_{\text{air}} \right] - \frac{\exp \left[\frac{10160}{R^*T} \right] \ln \left[\frac{\rho_i - \rho_0}{\rho_i - 550} \right]}{11A} \right) \right\} \right. \quad (3)$$

where R^* is the universal gas constant ($8.314 \text{ J mol}^{-1} \text{ K}^{-1}$), ρ_i is the density of ice (917 kg m^{-3}) and τ_{air} is the age of air in firn at lock-in, 15 yr (calculated based on ref. 17).

An important assumption for this approach to be valid is that the densification process was at steady state. We take the constancy of $\delta^{18}\text{O}_{\text{ice}}$ and accumulation⁹ during the Younger Dryas period as an indication that this assumption is warranted. Uncertainties are propagated assuming that they are uncorrelated and normally distributed via a Monte Carlo method. The largest source of uncertainty in the result arises from poor knowledge of the lock-in density ρ_L . Schwander²⁵ estimated this value at $795\text{--}800 \text{ kg m}^{-3}$, although at very low accumulation rates it may increase due to breakup of annual layers by wind erosion²⁵. We adopt a value of 820 kg m^{-3} based on an empirical formula^{17,45} for the variation of ρ_L with temperature, and note that the present value at the South Pole inferred from firn air $\delta^{15}\text{N}$ is $820 \pm 5 \text{ kg m}^{-3}$ (ref. 42; M. Battle, personal communication) and conditions (A and T) there are similar to

those inferred for the GISP2 site during Younger Dryas time⁸. We note that Greenland summit was not among the 17 sites used to develop the Herron–Langway model, so there is no circularity in our use of the GISP2 site as test of the method's ability to predict modern temperature.

Received 16 April; accepted 10 November 1997.

1. Johnsen, S. J. *et al.* Irregular glacial interstadials recorded in a new Greenland ice core. *Nature* **359**, 311–313 (1992).
2. Stuiver, M., Grootes, P. M. & Braziunas, T. F. The GISP2 $\delta^{18}\text{O}$ climate record of the past 16,500 years and the role of the sun, ocean, and volcanoes. *Quat. Res.* **44**, 341–354 (1995).
3. Alley, R. B. *et al.* Abrupt increase in Greenland snow accumulation at the end of the Younger Dryas event. *Nature* **362**, 527–529 (1993).
4. Chappellaz, J. *et al.* Synchronous changes in atmospheric CH_4 and Greenland climate between 40 and 8 kyr BP. *Nature* **366**, 443–445 (1993).
5. Brook, E. J., Sowers, T. & Orchard, J. Rapid variations in atmospheric methane concentration during the past 110,000 years. *Science* **273**, 1087–1093 (1996).
6. Chappellaz, J., Barnola, J.-M., Raymond, D., Korotkevich, Y. S. & Lorius, C. Ice-core record of atmospheric methane over the past 160,000 years. *Nature* **345**, 127–131 (1990).
7. Chappellaz, J. *et al.* Changes in the atmospheric CH_4 gradient between Greenland and Antarctica during the Holocene. *J. Geophys. Res.* **102**, 15987–15997 (1997).
8. Cuffey, K. M. *et al.* Large Arctic temperature change at the Wisconsin-Holocene glacial transition. *Science* **270**, 455–458 (1995).
9. Cuffey, K. M. & Clow, G. D. Temperature, accumulation, and ice sheet elevation in central Greenland through the last deglacial transition. *J. Geophys. Res.* **102**, 26383–26396 (1997).
10. Jouzel, J. *et al.* Validity of the temperature reconstruction from water isotopes in ice cores. *J. Geophys. Res.* **102**, 26471–26487 (1997).
11. Charles, C., Rind, D., Jouzel, J., Koster, R. D. & Fairbanks, R. G. Glacial-interglacial changes in moisture sources for Greenland: Influences on the ice core record of climate. *Science* **263**, 508–511 (1994).
12. Schwander, J., Stauffer, B. & Sigg, A. Air mixing in firn and the age of the air at pore close-off. *Ann. Glaciol.* **10**, 141–145 (1988).
13. Schwander, J. *et al.* The age of the air in the firn and the ice at summit, Greenland. *J. Geophys. Res.* **98**, 2831–2838 (1993).
14. Sowers, T., Bender, M. & Raynaud, D. Elemental and isotopic composition of occluded O_2 and N_2 in polar ice. *J. Geophys. Res.* **94**, 5137–5150 (1989).
15. Meese, D. A. *et al.* The Greenland Ice Sheet Project 2 depth-age scale: Methods and results. *J. Geophys. Res.* **102**, 26411–26423 (1997).
16. Herron, M. M. & Langway, C. C. Firn densification: An empirical model. *J. Glaciol.* **25**, 373–385 (1980).
17. Schwander, J. *et al.* Age scale of the air in the summit ice: Implication for glacial-interglacial temperature change. *J. Geophys. Res.* **102**, 19483–19494 (1997).
18. Grew, K. E. & Ibbs, T. L. *Thermal Diffusion in Gases* (Cambridge Univ. Press, 1952).
19. Chapman, S. & Cowling, T. G. *The Mathematical Theory of Non-Uniform Gases* (Cambridge Univ. Press, 1970).
20. Chapman, S. & Dootson, F. W. A note on thermal diffusion. *Phil. Mag.* **33**, 248–253 (1917).
21. Mariotti, A. Atmospheric nitrogen is a reliable standard for natural ^{15}N abundance measurements. *Nature* **303**, 685–687 (1983).
22. Dalton, J. On the constitution of the atmosphere. *Phil. Trans. Part 1*, 174–187 (1826).
23. Gibbs, J. W. *Collected Works Vol. 1, Thermodynamics* (Yale Univ. Press, New Haven, 1928).
24. Craig, H., Horibe, Y. & Sowers, T. Gravitational separation of gases and isotopes in polar ice caps. *Science* **242**, 1675–1678 (1988).
25. Schwander, J. in *The Environmental Record in Glaciers and Ice Sheets* (eds Oeschger, H. & Langway, C.)

53–67 (Wiley, New York, 1989).

26. Severinghaus, J. P., Bender, M. L., Keeling, R. F. & Broecker, W. S. Fractionation of soil gases by diffusion of water vapor, gravitational settling, and thermal diffusion. *Geochim. Cosmochim. Acta* **60**, 1005–1018 (1996).
27. Severinghaus, J. P. & Sowers, T. Thermal diffusion as a temperature-change indicator in ice core climate records. (abstr.) *Eos* **76**, F291 (1995).
28. Paterson, W. S. B. *The Physics of Glaciers* (Pergamon, Oxford, 1969).
29. Mayewski, P. A. *et al.* The atmosphere during the Younger Dryas. *Science* **261**, 195–197 (1993).
30. Alley, R. B. *et al.* Changes in continental and sea-salt atmospheric loadings in central Greenland through the most recent deglaciation. *J. Glaciol.* **41**, 503–514 (1995).
31. Sowers, T., Bender, M., Raynaud, D. & Korotkevich, Y. S. $\delta^{15}\text{N}$ of N_2 in air trapped in polar ice: A tracer of gas transport in the firn and a possible constraint on ice age–gas age differences. *J. Geophys. Res.* **97**, 15683–15697 (1992).
32. Colbeck, S. C. Air movement in snow due to windpumping. *J. Glaciol.* **35**, 209–213 (1989).
33. Johnsen, S. J., Dansgaard, W. & White, J. W. C. The origin of Arctic precipitation under present and glacial conditions. *Tellus B* **41**, 452–468 (1989).
34. Fairbanks, R. G. The age and origin of the “Younger Dryas climate event” in Greenland ice cores. *Paleoceanography* **5**, 937–948 (1990).
35. Dansgaard, W., White, J. W. C. & Johnsen, S. J. The abrupt termination of the Younger Dryas climate event. *Nature* **339**, 532–534 (1989).
36. Chappellaz, J., Fung, J. Y. & Thompson, A. M. The atmospheric CH_4 increase since the Last Glacial Maximum. *Tellus B* **45**, 242–257 (1993).
37. Martinie, P., Brasseur, G. P. & Granier, C. The chemical composition of ancient atmospheres: A model study constrained by ice core data. *J. Geophys. Res.* **100**, 14291–14304 (1995).
38. Prinn, R. G. *et al.* Atmospheric trends and lifetime of CH_4 , CCl_4 and global OH concentrations. *Science* **269**, 187–192 (1995).
39. Duplessy, J. C. *et al.* Changes in surface salinity of the North Atlantic Ocean during the last deglaciation. *Nature* **358**, 485–488 (1992).
40. Broecker, W. S. & Denton, G. H. The role of ocean-atmosphere reorganizations in glacial cycles. *Geochim. Cosmochim. Acta* **53**, 2465–2501 (1989).
41. Bender, M. L., Sowers, T., Barnola, J.-M. & Chappellaz, J. Changes in the O_2/N_2 ratio of the atmosphere during recent decades reflected in the composition of air in the firn at Vostok Station, Antarctica. *Geophys. Res. Lett.* **21**, 189–192 (1994).
42. Battle, M. *et al.* Atmospheric gas concentrations over the past century measured in air from firn at the South Pole. *Nature* **383**, 231–235 (1996).
43. Alley, R. B. & Koci, B. R. Recent warming in central Greenland? *Ann. Glaciol.* **14**, 6–8 (1990).
44. Taylor, K. C. *et al.* The ‘flickering switch’ of late Pleistocene climate change. *Nature* **361**, 432–436 (1993).
45. Martinie, P., Raynaud, D., Etheridge, D. M., Barnola, J.-M. & Mazaudier, D. Physical and climatic parameters which influence the air content in polar ice. *Earth Planet. Sci. Lett.* **112**, 1–13 (1992).
46. Etheridge, D. M. *et al.* Natural and anthropogenic changes in atmospheric CO_2 over the last 1000 years from air in Antarctic ice and firn. *J. Geophys. Res.* **101**, 4115–4128 (1996).
47. Bender, M. *et al.* Climate correlations between Greenland and Antarctica during the past 100,000 years. *Nature* **372**, 663–666 (1994).

Acknowledgements. We thank W. Broecker for discussions; R. Keeling for prompting our study of thermal diffusion; J. Schwander and R. Francey for reviews; B. Luz for developing the method for analysis of Ar isotopes in air; M. Swanson for the CH_4 analyses; J. Orchard for laboratory assistance; K. Cuffey, G. Clow and J. Schwander for providing pre-publication manuscripts; and the staff of the National Ice Core Laboratory for assistance in ice handling. J.P.S. was supported by a NOAA Climate and Global Change postdoctoral fellowship. We thank US NSF and US DOE (National Institutes of Global Environmental Change) for grant support.

Correspondence and requests for materials should be addressed to J.P.S. (e-mail: jeffs@gsosun1.gso.uri.edu).

KNOW YOUR COPY RIGHTS RESPECT OURS

The publication you are reading is protected by copyright law. Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.

Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931 Fax: 0171 436 3986

An exceptionally well-preserved theropod dinosaur from the Yixian Formation of China

Pei-ji Chen*, Zhi-ming Dong† & Shuo-nan Zhen‡

* Nanjing Institute of Geology and Palaeontology, Academia Sinica, 39 East Beijing Road, Nanjing 210008, People's Republic of China

† Institute of Vertebrate Paleontology and Paleoanthropology, Academia Sinica, PO Box 643, Beijing 100044, People's Republic of China

‡ Beijing Natural History Museum, 126 Tien Qiao Street, Beijing 100050, People's Republic of China

Two spectacular fossilized dinosaur skeletons were recently discovered in Liaoning in northeastern China. Here we describe the two nearly complete skeletons of a small theropod that represent a species closely related to *Compsognathus*. *Sinosauropteryx* has the longest tail of any known theropod, and a three-fingered hand dominated by the first finger, which is longer and thicker than either of the bones of the forearm. Both specimens have interesting integumentary structures that could provide information about the origin of feathers. The larger individual also has stomach contents, and a pair of eggs in the abdomen.

The Jehol biota¹ was widely distributed in eastern Asia during latest Jurassic and Early Cretaceous times. These freshwater and terrestrial fossils include macroplants, palynomorphs, charophytes, flagellates, conchostracans, ostracods, shrimps, insects, bivalves, gastropods, fish, turtles, lizards, pterosaurs, crocodiles, dinosaurs, birds and mammals. In recent years, the Jehol biota has become famous as an abundant source of remains of early birds^{2,3}. Dinosaurs are less common in the lacustrine beds, but the specimens described here consist of two nearly complete skeletons of a small theropod discovered by farmers in Liaoning. The skeletons are from the basal part of the Yixian Formation, from the same horizon as the fossil birds *Confuciusornis* and *Liaoningornis*³. Both are remarkably well preserved, and include fossilized integument, organ pigmentation and abdominal contents. One of the two was split into part and counterpart, and the sections were deposited in two different institutions. One side (in the National Geological Museum of China, Beijing) became the holotype of *Sinosauropteryx prima*, a supposed bird⁴. The counterpart and the second larger specimen are in the collections of the Nanjing Institute of Geology and Palaeontology.

The Yixian Formation is mainly composed of andesites, andesite-

breccia, agglomerates and basalts, but has four fossil-bearing sedimentary intercalations that are rich in tuffaceous materials. The Jianshangou (formerly Jianshan^{5,6}) intercalated bed (60 m thick) is the basal part of this volcanic sedimentary formation, and is made up of greyish-white, greyish-yellow and greyish-black sandstones, siltstones, mudstones and shales. These sediments are rich in fossils of mixed Jurassic-Cretaceous character. The primitive nature of the fossil birds of the Jianshangou fossil group has led to suggestions that the beds could be as early as Tithonian in age². But although *Confuciusornis* and the other birds³ are more advanced than *Archaeopteryx* in a number of significant features, we can only conclude that the beds that the fossils came from are probably younger than the Solnhofen Lithographic Limestones (Early Tithonian). The presence of *Psittacosaurus* in the same beds is more consistent with an Early Cretaceous age⁷, as are the palynomorphs⁸ and a recent radiometric date of the formation⁹, but other radiometric dating attempts have indicated older ages¹⁰.

Dinosauria OWEN 1842

Theropoda MARSH 1881

Coelurosauria VON HUENE 1914

Compsognathidae MARSH 1882

Sinosauropteryx prima JI and JI 1996

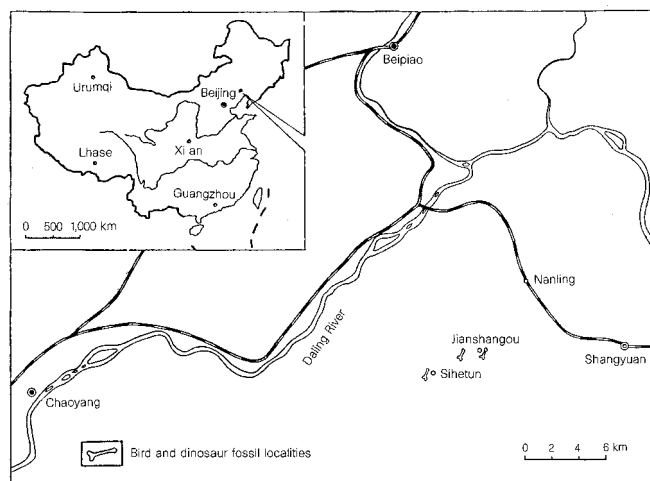


Figure 1 Map showing the localities of *Sinosauropteryx* and *Confuciusornis* in western Liaoning, northeastern China.

Holotype. Part (National Geological Museum of China, GMV 2123) and counterpart (Nanjing Institute of Geology and Palaeontology NIGP 127586) slabs of a complete skeleton.

Referred specimen. Nanjing Institute of Geology and Palaeontology NIGP 127587. Nearly complete skeleton, lacking only the distal half of the tail.

Locality and horizon. Jianshangou-Sihetun area of Beipiao, Liaoning, People's Republic of China. Yixian Formation, Jehol Group, Upper Jurassic or Lower Cretaceous (Fig. 1).

Diagnosis. Compsognathid with longest tail known for any theropod (64 caudals). Skull 15% longer than femur, and forelimb (humerus plus radius) only 30% length of leg (femur plus tibia), in contrast with *Compsognathus* where skull is same length as femur, and forelimb length is 40% leg length. Within the Compsognathidae, forelimb length (compared to femur length) is shorter in *Sinosauropteryx* (61–65%) than it is in *Compsognathus* (90–99%). In contrast with all other theropods, ungual phalanx II–2 is longer than the radius. Haemal spines simple and spatulate, whereas those of *Compsognathus* taper distally.

Description

Sinosauropteryx is comparable in size and morphology to known specimens of *Compsognathus*^{11,12} from Germany and France. The smaller Chinese specimen (Fig. 2) is 0.68 m long (snout to end of tail) and has a femur length of 53.2 mm, whereas the second specimen (Fig. 3) has a femur length of 86.4 mm. The former is smaller than the type specimen of *Compsognathus longipes* (femur length about 67 mm) and the latter is smaller than the second specimen of *Compsognathus* from Canjuers (France), which has a femur length of 110 mm and an estimated length of 1.25 m. Although size and body proportions indicate that the smaller specimen was younger when it died, well-ossified limb joints and tarsals suggest that it was approaching maturity.

Sinosauropteryx and *Compsognathus* share several characteristics that indicate close relationship. These can be used to diagnose the Compsognathidae and include unserrated premaxillary but serrated maxillary teeth, a powerful manual phalanx I–1 (shaft diameter is

greater than that of the radius), fan-shaped neural spines on the dorsal vertebrae, limited anterior expansion of the pubic boot and a prominent obturator process of the ischium.

Other characteristics were used to diagnose *Compsognathus*, including the presence of a relatively large skull and short forelimbs. In *Compsognathus*, skull length is 30% of that of the presacral vertebral column, whereas this same ratio is 40% in the new specimen NIGP 127586 and 36% in NIGP 127587. Unfortunately, relative skull length is highly variable in theropods. Comparing skull length with femur length, which is less variable than vertebral length, most theropods have skulls 100–119% the length of the femur¹³. The *Compsognathus* skulls are 99–100% and *Sinosauropteryx* skulls are 113–117%. Compsognathids have short forelimbs¹¹, 40% of the length of the hindlimb in *Compsognathus*. In *Sinosauropteryx*, the lengths of humerus plus radius divided by the sum of femur and tibia lengths produces a figure of less than 30% (Table 1). Unfortunately, such ratios are dependent on the absolute

Table 1 Comparison of size and proportions of *Sinosauropteryx* and *Compsognathus*

Species	Specimen	Skull	Humerus	Radius	Femur	Tibia	Skull/femur	Arm/leg
<i>Compsognathus</i> sp.	BSP ASI	70	39	24.7	71	87.6	0.99	0.40
<i>Compsognathus</i> sp.	MNHN	110	67	42	110		1.00	
<i>Sinosauropteryx primus</i>	NIGP 127586	62.5	20.3	12.4	53.2	61	1.17	0.29
<i>Sinosauropteryx primus</i>	NIGP 127587	97.2	35.5	21	86.4	97	1.13	0.31

Length measurements are given in millimetres. Data about *Compsognathus* are from ref. 11.

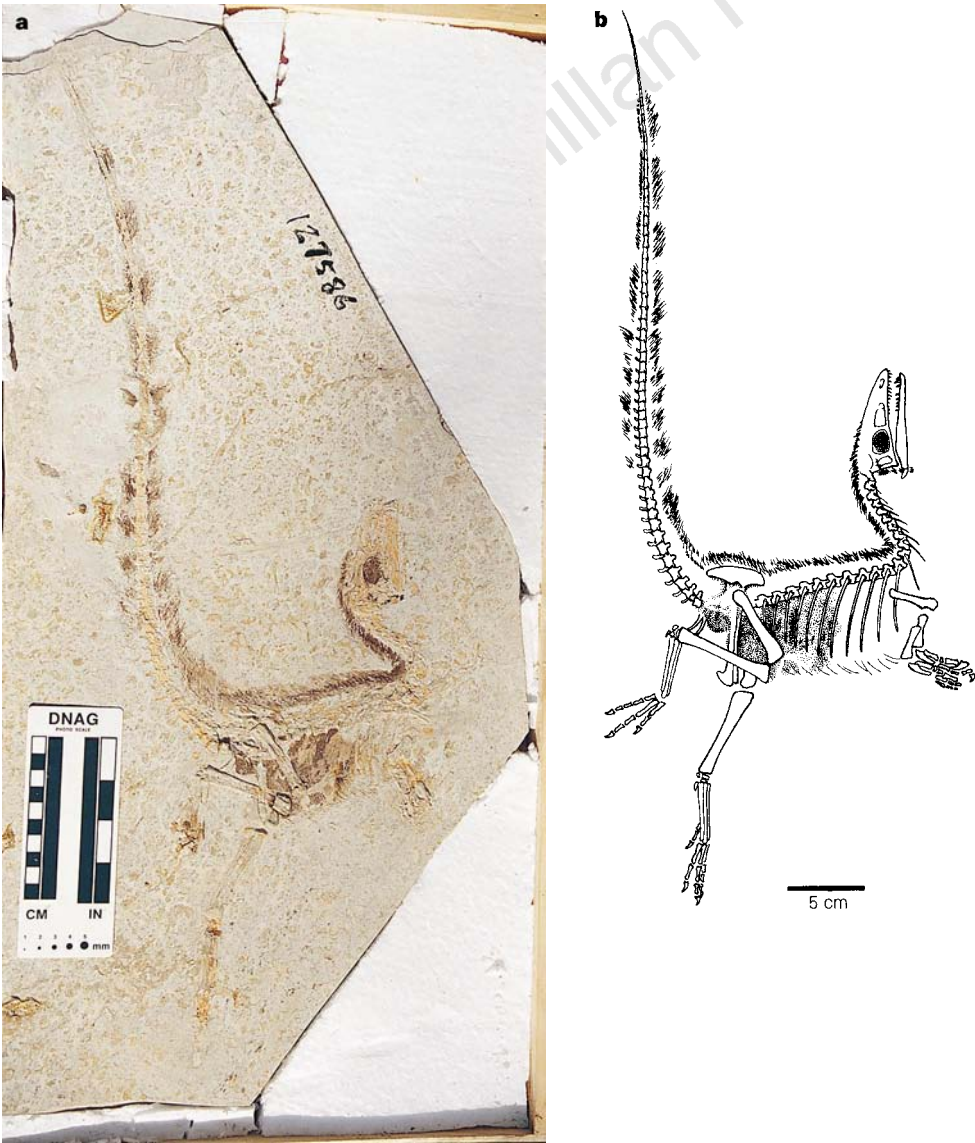


Figure 2 *Sinosauropteryx prima* Ji and Ji. **a**, NIGP 127586, the counterpart of holotype (GMV 2123). **b**, Skeletal reconstruction of NIGP 127586. The integumentary structures are along the dorsal side and tail and dark pigmentation in the abdominal region might be some soft tissues of viscera.

size of the animal, mostly because of negative allometry experienced by the tibia during growth or interspecific size increase. Comparing the lengths of humerus plus radius with femur length produces more useful results. The resulting figures fall within the range of most theropods (60–110%). The abelisaurid *Carnosaurus* and all tyrannosaurs have relatively shorter arms. Within the Compsognathidae, however, arm length is shorter in *Sinosauropteryx* (61–65%) than it is in *Compsognathus* (90–99%).

Both specimens of *Sinosauropteryx* have 10 cervical and 13 dorsal vertebrae. The posterior cervical vertebrae have biconcave centra. We could not determine the number of sacra. The tail is extremely long. In the smaller specimen it is almost double the snout–vent length, and there are 59 caudal vertebrae exposed with an estimated five more than have been lost from the middle of the tail of NIGP 127586 (but present in GMV 2123). Only the first 23 vertebrae are preserved in the larger specimen, but this section is longer than the summed lengths of the cervical, dorsal and sacral vertebrae. Neither of the European specimens has a complete tail, but in both cases the tail was clearly longer than the body. When vertebral lengths are normalized (divided by the average lengths of caudal vertebrae 2–5), there are no significant differences between vertebral lengths in any of the four tails. As in *Compsognathus*, the dorsal neural spines are peculiar in that they are anteroposteriorly long but low, and often are fan-shaped.

The caudal centra increase in length over the first six segments, but posteriorly decrease progressively in length and all other dimensions. The first 10 tail vertebrae have neural spines, most of which slope posterodorsally. There are at least four pairs of caudal ribs in NIGP 127586, and more distal caudals have low bumps in this region that could also be interpreted as transverse processes. This could be another way to distinguish the Asian and European compsognathids, because the German specimen of *Compsognathus* apparently lacks caudal ribs and transverse processes¹¹. Haemal spines are found on at least the first 47 caudals of NIGP 127586, and the anterior ones are simple spatulate structures that curve gently posteroventrally. The haemal spines are oriented more posteriorly than ventrally, and are more strongly curved.

Both specimens have 13 pairs of dorsal ribs. The ribs indicate a high but narrow body. The distal ends of the first two pairs of ribs are expanded and end in cup-like depressions that suggest the presence of a cartilaginous sternum. The gastralia are well preserved with two gastralia on each side of a segment. The median gastralia cross to form the interconnected ‘zig-zag’ pattern characteristic of

all theropods¹⁴ and primitive birds like *Archaeopteryx* and *Confuciusornis*.

The front limb is relatively short and stout. Both NIGP 127586 and NIGP 127587 have articulated hands, something that is lacking in the two European specimens. What has been interpreted by some as the first metacarpal¹¹ in *Compsognathus* is the first phalanx of digit I, as was originally proposed by von Huene¹⁵. The first metacarpal is short (4.2 mm long in NIGP 127586, and double that length in 127587), and is probably the element identified as a carpal in the French specimen¹². As is typical of all theropods, the collateral ligament pits of the first phalanx are much closer to the extensor surface of the bone than they are to the flexor surface. Both phalanx I–1 and the ungual that it supports are relatively large, each being as long as the radius, and thicker than the shafts and the distal ends of either the radius or the ulna. This unusual character seems to have been partially developed in at least phalanx I–1 of *Compsognathus*. Relative to the length of the radius, both these elements are longer in *Sinosauropteryx* than in any other known theropod except for *Mononykus*¹⁶. As indicated by the proposed phylogenetic placement¹⁶ of *Mononykus*, there are too many anatomical differences between compsognathids and *Mononykus* to suggest a close relationship, and the similarities probably represent convergence.

The long (39 mm in NIGP 127586, 67.5 mm in NIGP 127587), low (22.2 mm high at both pubic and ischial peduncles in NIGP 127587) ilium is shallowly convex on the dorsal side in lateral aspect. The pubis, which is 82.8 mm long in NIGP 127587, is oriented anteroventrally, but is closer to vertical than it is in most non-avian theropods. The distal end expands into a pubic boot as in most tetanuran theropods. In the larger specimen, this expansion is 17.7 mm. As in *Archaeopteryx*¹⁷, *Compsognathus*¹¹ and dromaeosaurs¹⁸, the boot expands posteriorly from the shaft of the pubis, and the anterior expansion is moderate. The lack of the significant anterior expansion of the pubic boot may be correlated with the inclination of the shaft. The ischium is only two-thirds the length of the pubis in NIGP 127587. It tapers distally into a narrow shaft (3.2 mm in diameter), and like *Compsognathus*, there is a slight expansion at the end (6 mm in NIGP 127587). The prominent obturator process is also found in *Compsognathus*.

The shaft of the femur is gently curved. Both tibia and fibula are elongate. The astragalus and calcaneum are present in both specimens, although not clearly seen in either. There are five metatarsals, but as in other theropods and early birds, the first is reduced to a distal articular condyle, and the fifth is reduced to a proximal splint (Fig. 2b). Metatarsals II, III and IV are closely appressed and elongate, but are not co-ossified. The second and fourth metatarsals do not contact each other. Pedal phalanges are conservative in number (2–3–4–5–0) and morphology.

Inclusions within the body cavity

Like the German *Compsognathus*, the larger Chinese specimen has stomach contents preserved within the rib cage. This consists of a semi-articulated skeleton of a lizard, complete with skull (Fig. 4a,b). Numerous lizard skeletons have been recovered from these beds, but have yet to be described. Low in the abdomen of NIGP 127587, anterior to and slightly above the pubic boot, lies a pair of small eggs (37 × 26 mm) (Fig. 4a,c), one in front of the other. Additional eggs may lie underneath. Gastralia lie over the exposed surfaces of the eggs, and the left femur protrudes from beneath them, so there can be no doubt that they were within the body cavity. It is possible that the eggs were eaten by the dinosaur. However, given their position in the abdomen behind and below the stomach contents, and the fact that they are in the wrong part of the body cavity for the egg shell to be intact, it is more likely that these were unlaidd eggs of the compsognathid. Eggs have also been reported in the holotype of *Compsognathus*¹⁹, but they are more numerous and are only 10 mm in diameter. As they were also found outside the body cavity, their

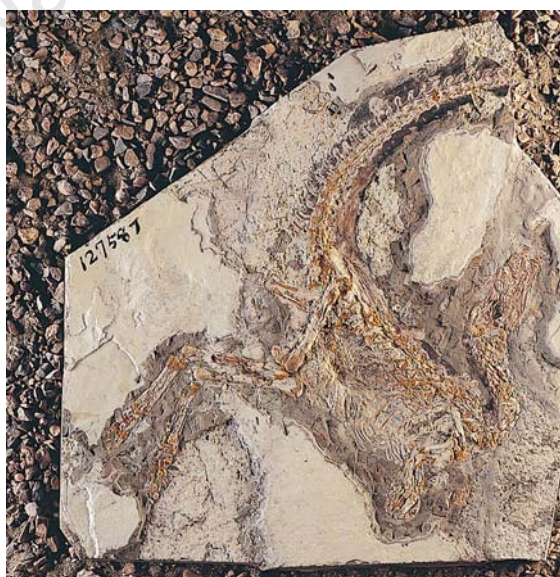


Figure 3 *Sinosauropteryx prima* Ji and Ji, NIGP 127587, an adult individual from the same locality as holotype.

identification as *Compsognathus* eggs has not been widely accepted. The presence of fewer but larger eggs in *Sinosauropteryx* casts additional doubt on this identification.

Although more than two eggs may have been present in the larger specimen of *Sinosauropteryx*, it does not seem as though many could have been held within the abdomen. It may well be that these dinosaurs laid fewer eggs than most (some species are known to have produced in excess of 40)²⁰. However, it is more likely that their presence demonstrates paired ovulation, as has been suggested for *Oviraptor*²¹, *Troodon*²² and other theropods. *Sinosauropteryx* therefore probably laid eggs in pairs, with a delay for ovulation between each pair.

Integumentary structures

One of the most remarkable features of both Chinese specimens is the preservation of integumentary structures. In the larger speci-

men, these structures can be seen along the dorsal surface of the neck and back, and along the upper and lower margins of the tail, but in the smaller specimen the integumentary structures are clearer (Fig. 5). They cover the top of the back half of the skull, the neck, the back, the hips and the tail. They also extend along the entire ventral margin of the tail. Small patches can be seen on the side of the skull (behind the quadrate and over the articular), behind the right humerus, and in front of the right ulna. With the exception of a small patch outside the left ribs of NIGP 127587 and several areas on the left side of the tail (lateral to the vertebrae), integumentary structures cannot be seen along the sides of the body. The structures were probably present in the living animals, as indicated by the density of the covering dorsal to the body, and by the few random patches of integumentary structures that can be seen elsewhere.

In the two theropods, the distances separating the integumentary structures from the underlying bones are directly proportional to the amount of skin and muscle that would have been present. As in modern animals, the integument closely adheres to the tops of the skull and hips, and becomes progressively closer to the caudal vertebrae towards the tip of the tail. In the posterior part of the neck, over the shoulders, and at the base of the tail, the integumentary structures are more distant from the underlying skeletal elements, and in life would have been separated by greater thicknesses of muscle and other soft tissues.

The orientation and frequently sinuous lines of the integumentary structures suggest they were soft and pliable, and semi-independent of each other. They frequently cross each other, and are tangled in some areas. There is an apparent tendency for the integumentary structures to clump along the tail of the smaller specimen, but this is an artefact of the splitting plane between NIGP 127586 (Fig. 2a) and GMV 2123. As both individuals were lying in the water of a lake when they were buried, it is clear that we are not looking at the normal orientation of the integumentary structures in the fossils. Under magnification, the margins of the larger structures are darker

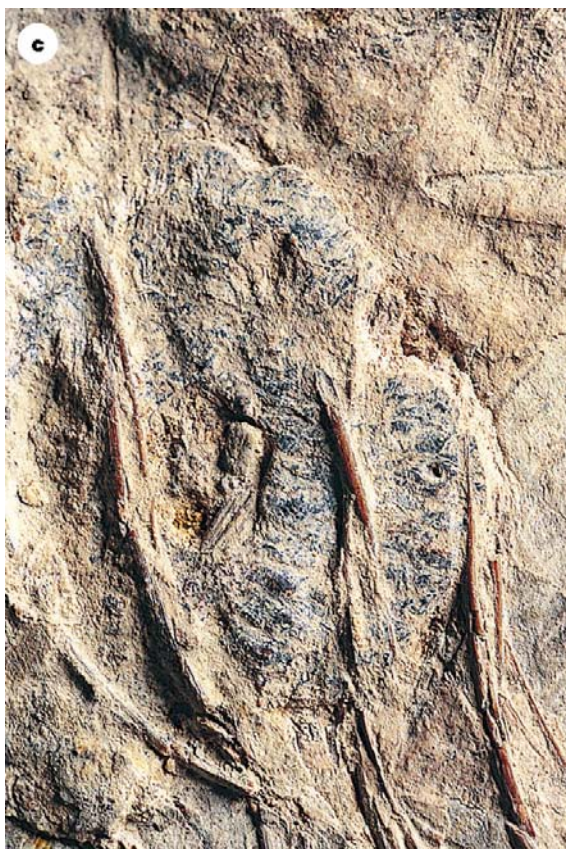


Figure 4 Body of NIGP 127587. **a**, Stomach contents are preserved within the rib cage, and include a small lizard and a pair of eggs. **b**, A close-up of the lizard skull. **c**, A close-up of a pair of the eggs.

along the edges, but lighter medially, which indicates that they might have been hollow. Overall, the integumentary structures are rather coarse for such a small animal, and the thickest strands are much thicker than the hairs of the vast majority of small mammals²³. In NIGP 127586, integumentary structures are first seen on the dorsal surface of the skull in front of the orbit. The skull is semidisarticulated, and sediment still covers the snout region, so it is possible that the integumentary structures extended more anteriorly. The most rostral integumentary structures are 5.5 mm long, and extend about 4 mm above the skull. They quickly lengthen to at least 21 mm above the distal ends of the scapulae. This axial length seems to stay constant along most of the back, but decreases sharply to 16 mm dorsal to the ilium. The longest integumentary structures seem to have been above the base of the tail, although it is impossible to measure any single structure. More distally along the tail, integumentary structures decrease more rapidly on the lower side of the tail than on the upper. By caudal 47, the ventral structures are 4.2 mm long, about half the length of the dorsal structures in that region.

The size distribution of the integumentary structures of NIGP 127587 follow the same general pattern as in the smaller specimen. Although the integument tends to look thinner on this specimen, it is simply because the integumentary structures are lying closer to the body. Individual measurements are consistently larger than those of NIGP 127586. The integumentary structures are 13 mm long above the skull, 23.5 mm above the fourth cervical, at least 35 mm over the scapulae, at least 40 mm over caudal 27 (Fig. 6), and at least 35 mm below caudal 25. Integumentary structures on the left side of the body are largely covered by ribs, gastralia, stomach contents and matrix, so it is only possible to say that each is more than 5 mm long. Those associated with the right ulna are 14 mm long.

Integumentary structures have also been reported in the theropod *Pelecanimimus*²⁴ from the Lower Cretaceous of Spain. These consist of subparallel fibres arranged perpendicular to the bones, with a less conspicuous secondary system parallel to them. As described, they seem to be similar to the integumentary structures of *Sinosauropteryx*.

Skin impressions have been found on most main types of dinosaurs, including sauropods, ankylosaurs, ornithomorphs, stego-

saurs, ceratopsians, and several genera of large theropods. In all of these animals, there is no evidence of integumentary structures, and the skin usually has a 'pebbly' surface texture. Integumentary structures have been claimed for both specimens of *Compsognathus*^{11,25} though the interpretations have been questioned in both cases¹¹. In the German specimen, there was supposedly a patch of skin over the abdominal region. The French specimen included some strange markings in the region of the forearm, that were originally identified as a swimming appendage formed either of dermal bone or of thick skin¹², but it is clearly not well enough preserved to be positively identified. The identification of these structures as integumentary is questionable¹¹, and there is nothing on the Chinese specimens to support the presence of such structures in compsognathids. Evidence of feathers in *Compsognathus* was sought¹¹ without success, but this lack of evidence on the German specimen of *Compsognathus* does not eliminate the possibility that they might have existed.

Discussion

The integumentary structures of *Sinosauropteryx* are extremely interesting regardless of whether they are referred to as feathers, protofeathers, or some other structure. Unfortunately, they are piled so thick that we have been unable to isolate a single one for examination. Comparison with birds from the same locality shows that the same problem exists with identifying individual feathers (other than the flight feathers) and components of feathers in avian specimens. The morphological characteristics that we describe suggest that the integumentary structures seem to resemble most closely the plumules of modern birds, having relatively short quills and long, filamentous barbs. The absence of barbules and hooklets is uncommon in modern birds, but has been noted in Cretaceous specimens²⁶.

It has been proposed that the feathers of another recently discovered animal from the same locality in Liaoning are structurally intermediate between the integumentary structures of *Sinosauropteryx* and the feathers of *Archaeopteryx*²⁷. The clearly preserved feathers of *Protarchaeopteryx robusta* are symmetrical, which indicates that the animal was not capable of flight. This is confirmed by the relatively short length of the forelimb. Both *Sinosauropteryx* and *Protarchaeopteryx* had been identified as



Figure 5 (Far left) Integumentary structures in the neck and dorsal sides of NIGP 127586.

Figure 6 (Left) Integumentary structures over caudal 27 of NIGP 127587.

birds because of the presence of feathers²⁷, but much more work needs to be done to prove that the integumentary structures of *Sinosauropteryx* have any structural relationship to feathers, and phylogenetic analysis of the skeleton clearly places compsognathids far from the ancestry of birds. Despite arguments to the contrary²⁸, cladistic analysis favours the notion that the bird lineage originated within theropod dinosaurs^{29,30}. If this phylogenetic framework is accepted, the integumentary structures of *Sinosauropteryx* could shed light on some of the many hypotheses concerning feather origins. Three main functions have been suggested for the initial development of feathers—display, aerodynamics and insulation.

The integumentary structures of *Sinosauropteryx* have no apparent aerodynamic characteristics, but might be representative of what covered the ancestral stock of birds. It is highly unlikely that something as complex as a bird feather could evolve in one step, and many animals glide and fly with much simpler structures. Even birds secondarily simplify feathers when airborne flight ceases to be their main method of locomotion, and produce structures that are intermediate between reptilian scales and feathers³¹. The multi-branched integumentary structures of the Chinese compsognathids are relatively simple, but are suitable for modification into the more complex structures required for flight.

Feathers may have appeared first as display structures³¹, but the density, distribution, and relatively short lengths of the integumentary structures of *Sinosauropteryx* suggest that they were not used for display. It is conceivable that both specimens are female, and that the males had more elaborate integumentary structures for display. It is also possible that the integumentary structures were coloured to serve a display function. Therefore, the existing *Sinosauropteryx* specimens do not support the hypothesis that feathers evolved primarily for display, but do not disprove it either.

The dense, pliable integumentary structures of the Chinese compsognathids would not have been appropriate as heat shields to screen and shade the body from the Sun's rays³². Although they may have been effective in protecting the body from solar radiation in warm weather, they would also have been effective in preventing an ectothermic theropod from rapidly warming up by basking in the sunshine. If small theropods were endothermic, they would have needed insulation to maintain high body temperatures^{33–35}. The presence of dense integumentary structures may suggest that *Sinosauropteryx* was endothermic, and that heat retention was the primary function for the evolution of integumentary structures^{36–38}. Recently published histological studies suggest at least some early birds were not truly endothermic³⁹, although they may have been physiologically intermediate between poikilothermic ectotherms and homeothermic endotherms⁴⁰.

The Chinese compsognathids have integumentary structures consisting of vertical fibres running from the base of the head along the back and around the tail extending forwards almost to the legs. There are no structures showing the fundamental morphological features of modern bird feathers, but they could be previously unidentified protofeathers which are not as complex as either down feathers or even the hair-like feathers of secondarily flightless birds. Their simplicity would not have made them ineffective for insulation when wet any more than it negates the insulatory capabilities of mammalian hair. We cannot determine whether or not the integumentary structures were arranged in pterygiae, but they were long enough to cover apteria, if they existed, and could therefore still have been effective in thermoregulation. Continuous distribution is not essential to be effective in this function²⁸, especially if the apteria are part of a mechanism for dispersing excess heat. Finally, the aerodynamic capabilities of bird feathers are not comprised by the previous evolution of less complex protofeathers that had some other function, such as insulation.

In addition to the integumentary structures, there is dark pigmentation over the eyes of both specimens. A second region of dark pigmentation in the abdominal region of the smaller specimen

might represent some soft tissues of viscera.

Multidisciplinary and multinational research is just beginning on these unique small theropods. Techniques developed to study fossil feathers^{38,41} will be useful research tools as work progresses. In the meantime, the integumentary structures of *Sinosauropteryx* suggest that feathers evolved from simpler, branched structures that evolved in non-avian theropod dinosaurs, possibly for insulation. □

Received 14 January; accepted 18 September 1997.

- Chen, P. J. Distribution and migration of Jehol fauna with reference to nonmarine Jurassic–Cretaceous boundary in China. *Acta Palaeontol. Sin.* **27**, 659–683 (1988).
- Hou, L.-H., Zhang, J.-Y., Martin, L. D. & Feduccia, A. A beaked bird from the Jurassic of China. *Nature* **377**, 616–618 (1995).
- Hou, L.-H., Martin, L. D., Zhang, J.-Y. & Feduccia, A. Early adaptive radiation of birds: evidence from fossils from northeastern China. *Science* **274**, 1164–1167 (1996).
- Ji, Q. & Ji, S. A. On discovery of the earliest bird fossil in China and the origin of birds. *Chinese Geol.* **233**, 30–33 (1996).
- Chen, P. J. et al. Studies on the Late Mesozoic continental formations of western Liaoning. *Bull. Nanjing Inst. Geol. Palaeontol.* **1**, 22–25 (1980).
- Chen, P. J. Nonmarine Jurassic strata of China. *Bull. Mus. N. Arizona* **60**, 395–412 (1996).
- Dong, Z. M. Early Cretaceous dinosaur faunas in China: an introduction. *Can. J. Earth Sci.* **30**, 2096–2100 (1993).
- Li, W. B. & Liu, Z. S. The Cretaceous palynofloras and their bearing on stratigraphic correlation in China. *Cretaceous Res.* **15**, 333–365 (1994).
- Smith, P. E. et al. Dates and rates in ancient lakes: ⁴⁰Ar–³⁹Ar evidence for an Early Cretaceous age for the Jehol Group, northeast China. *Can. J. Earth Sci.* **32**, 1426–1431 (1995).
- Wang, D. F. & Diao, N. C. Geochronology of Jura–Cretaceous volcanics in west Liaoning, China. *Scientific papers on geology for international exchange* **5**, 1–12 (Geological Publishing House, Beijing, 1984).
- Ostrom, J. H. The osteology of *Compsognathus longipes* Wagner. *Zitteliana* **4**, 73–118 (1978).
- Bidar, A., Demay, L. & Thomel, G. *Compsognathus corallistris* nouvelle espèce de dinosaureien théropode du Portlandien de Canjuers (sud-est de la France). *Ann. Mus. d'Hist. Nat. Nice* **1**, 3–34 (1972).
- Currie, P. J. & Zhao, X. J. A new large theropod (Dinosauria, Theropoda) from the Jurassic of Xinjiang, People's Republic of China. *Can. J. Earth Sci.* **30**, 2037–2081 (1993).
- Claessens, L. Dinosaur gastralia and their function in respiration. *J. Vert. Palaeontol.* **16**, 28A (1996).
- von Huene, F. The carnivorous Saurischia in the Jura and Cretaceous formations principally in Europe. *Revista Museo Plata* **29**, 35–167 (1926).
- Perle, A., Chiappe, L. M., Barsbold, R., Clark, J. M. & Norell, M. Skeletal morphology of *Mononykus olecranus* (Theropoda: Avialae) from the Late Cretaceous of Mongolia. *Am. Mus. Novit.* **3105**, 1–29 (1994).
- Wellnhofer, P. Das siebte Exemplar von *Archaeopteryx* aus den Solnhofener Schichten. *Archaeopteryx* **11**, 1–48 (1993).
- Barsbold, R. Carnivorous dinosaurs from the Cretaceous of Mongolia. *Sovmestnaya Sovetskoye-Mongol'skaya Paleontol. Ekspiditsiya, Trudy* **19**, 5–119 (1983).
- Griffiths, P. The question of *Compsognathus* eggs. *Rev. Paleobiol. Spec. issue* **7**, 85–94 (1993).
- Carpenter, K., Hirsch, K. F. & Horner, J. R. *Dinosaur Eggs and Babies* (Cambridge Univ. Press, 1994).
- Dong, Z. M. & Currie, P. J. On the discovery of an oviraptorid skeleton on a nest of eggs at Bayan Mandahu, Inner Mongolia, People's Republic of China. *Can. J. Earth Sci.* **33**, 631–636 (1996).
- Varricchio, D. J., Jackson, E., Borkowski, J. J. & Horner, J. R. Nest and egg clutches of the dinosaur *Troodon formosus* and the evolution of avian reproductive traits. *Nature* **385**, 247–250 (1997).
- Meng, J. & Wyss, A. R. Multituberculate and other mammal hair recovered from Palaeogene excreta. *Nature* **385**, 712–714 (1997).
- Pérez-Moreno, B. P. et al. A unique multitoothed ornithomimid dinosaur from the Lower Cretaceous of Spain. *Nature* **370**, 363–367 (1994).
- von Huene, F. Der Vermuthliche Hautpanzer des *Compsognathus longipes* Wagner. *Neues Jb. F. Min.* **1**, 157–160 (1901).
- Grimaldi, D. & Case, G. R. A feather in amber from the Upper Cretaceous of New Jersey. *Am. Mus. Novit.* **3126**, 1–6 (1995).
- Ji, Q. & Ji, S. A. *Protarchaeopteryx*, a new genus of Archaeopterygidae in China. *Chinese Geol.* **238**, 38–41 (1997).
- Feduccia, A. *The Origin and Evolution of Birds* (Yale Univ. Press, New Haven, 1996).
- Gauthier, J. In *The Origin of Birds and the Evolution of Flight* (ed. Padian, K.) 1–55 (California Acad. Sci., San Francisco, 1986).
- Fastovsky, D. E. & Weishampel, D. B. *The Evolution and Extinction of the Dinosaurs* (Cambridge Univ. Press, 1996).
- McGowan, C. Feather structure in flightless birds and its bearing on the question of the origin of feathers. *J. Zool. (Lond.)* **218**, 537–547 (1989).
- Paul, G. S. *Predatory Dinosaurs of the World* (Simon and Schuster, New York, 1988).
- Ewart, J. C. The nestling feathers of the mallard, with observations on the composition, origin, and history of feathers. *Proc. Zool. Soc. Lond.* 609–642 (1921).
- Van Tyne, J. & Berger, A. J. *Fundamentals of Ornithology* (Wiley, New York, 1976).
- Young, J. Z. *The Life of Vertebrates* (Oxford Univ. Press, 1950).
- Chinsamy, A., Chiappe, L. M. & Dodson, P. Growth rings in Mesozoic birds. *Nature* **368**, 196–197 (1994).
- Chiappe, L. M. The first 85 million years of avian evolution. *Nature* **378**, 349–355 (1995).
- Brush, A. H. In *Avian Biology* vol. 9 (eds Farner, D. S., King, J. R. & Parkes, K. C.) 121–162 (Academic, London, 1993).
- Regal, P. J. The evolutionary origin of feathers. *Quart. Rev. Biol.* **50**, 35–66 (1975).
- Ostrom, J. H. Reply to 'Dinosaurs as reptiles'. *Evolution* **28**, 491–493 (1974).
- Davis, P. G. & Briggs, D. E. G. Fossilization of feathers. *Geology* **23**, 783–786 (1995).

Acknowledgements. This study was supported by NSFC. We thank L.-s. Chen and P. J. Currie (Royal Tyrrell Museum of Palaeontology) for helping to prepare the fossil materials and manuscript; M.-m. Zhang, X.-n. Mu, G. Sun, J. H. Ostram, A. Brush, L. Martin, P. Wellnhofer, N. J. Mateer, E. B. Koppelhus, D. B. Brinkman, D. A. Eberth, J. A. Ruben, L. Chiappe, S. Czerkas, R. O'Brien, D. Rimplinger, M. Vickaryous and D. Unwin for assistance and comments; and L. Mazzatenta and M. Skrepnick for help producing the photographs and drawings.

Correspondence and requests for materials should be addressed to P.-j.C. (e-mail: lpsnigp@nanjing.jpsta.chinamail.sprint.com).

A planetary companion for 51 Pegasi implied by absence of pulsations in the stellar spectra

David F. Gray

Department of Physics & Astronomy, University of Western Ontario, London, Ontario N6A 3K7, Canada

Systematic variations in the Doppler shifts of absorption lines in the spectrum of the star 51 Pegasi were interpreted as indicating the presence of a planet about half the mass of Jupiter, very close to the star^{1,2}. But that interpretation was called into question when variations in the line shapes that tracked the apparent orbital phase were reported^{3,4}; this suggested that a planet was an inadequate explanation of the radial-velocity data. Here I report results from recent monitoring of 51 Peg; the oscillations I previously published are not evident in the new data. When combined with two other high-precision observations of 51 Peg (refs 5–7), that also see no changes in line shape, a planet may indeed be the best explanation for the radial-velocity results.

High-resolution spectroscopic observations of 51 Peg were started at the University of Western Ontario in 1989, with a sample rate of a few exposures per year. (51 Peg is also known as HR8729 and HD217014, is of spectral type G5 V or IV, and has $B - V = 0.68$ and $V = 5.53$; where B and V are magnitudes in the B and V bands, respectively). When an unusual planet was claimed to be in orbit around this star^{1,2} (based on a 4.23-d periodic shift of the spectral lines), it seemed prudent to check my observations for this periodicity even though they were not spaced closely enough in time to independently discover such short periods. The fundamental problem one confronts with time spacing much larger than the period is severe aliasing, making it impossible to know which of the many aliases is the true one. But in the case of 51 Peg, the period was already identified from the line shifts, and so the question became simply, can this 4.23-d period be seen in my data? The parameters obtained from my spectroscopic data are a measure of the shape of the spectral lines, specifically (in the 51-Peg investigation) the slope and curvature of the line bisector of the Fe I line of wavelength 6,252.57 Å. Phasing these parameters with the 4.23-d period showed

a sinusoidal signal; the spectral lines appeared to be changing their shape on a 4.23-d cycle^{3,4}. And if this is true, then the orbit hypothesis, invoked to explain the radial-velocity-type shifts, becomes doubtful because although orbital motion can shift spectral lines in wavelength, it cannot alter their shape. The alternative explanation of non-radial oscillations on the surface of the star was suggested⁴. However, the reality of the shape variations was inconclusive because the amplitude of the signal was near the noise level and because the aliasing precluded a convincing identification of the true period. Nevertheless, the probability of noise producing this 4.23-d signal was significantly less than 1 in 300. It was therefore important to continue the observations of this star with increased frequency to resolve the issue.

New observations of 51 Peg were undertaken in several places, including the McDonald Observatory⁵, the Whipple Observatory^{6,7} and the University of Western Ontario as described here. My 1997 observations were taken with the same spectrograph and detector used to acquire the earlier data. A log of the 30 exposures is given in Table 1. The exposures span 101.78 d or just over 24 cycles of 4.23 d. Nominal signal-to-noise ratios in the continuum, based on the level of the recorded signal, range from 197 to 525, with a mean of only 336. This is about 20% worse than the earlier data, a result of poorer seeing and more cloud during the 1997 season compared to previous seasons.

The bisector curvature was computed by adding the velocity spans of the line bisectors^{3,8} between $F/F_c = 0.50$ and 0.70, and between 0.70 and 0.83, where F/F_c is the flux in the line profile normalized to the continuum. This is essentially the same parameter used in the Gray and Hatzes analysis⁴. The last column in Table 1 lists the bisector curvature for Fe I 6,252.57 Å. Although one could go through an exhaustive discussion of these new observations, it is sufficient for the purpose at hand to consider their periodogram (also called the Fourier transform or power spectrum) as given in Fig. 1. Figure 1a shows the window function corresponding to the time sample spacing of the data. A replication of this pattern will occur at the frequency of each Fourier component of signal in the data⁹. In Fig. 1b we show the data compared to the expected signal. Specially, we expected a peak at 0.2365 cycles d⁻¹ corresponding to the 4.23-d period found in the earlier investigation⁴, but there is no evidence of such a peak. More formally, the observed power at this frequency is seven times smaller than the expected power for a 45 m s⁻¹ oscillation (an amplitude consistent with the earlier data). The periodogram does not change significantly if the individual exposures are weighted according to the square root of their signal-to-noise ratios or left unweighted.

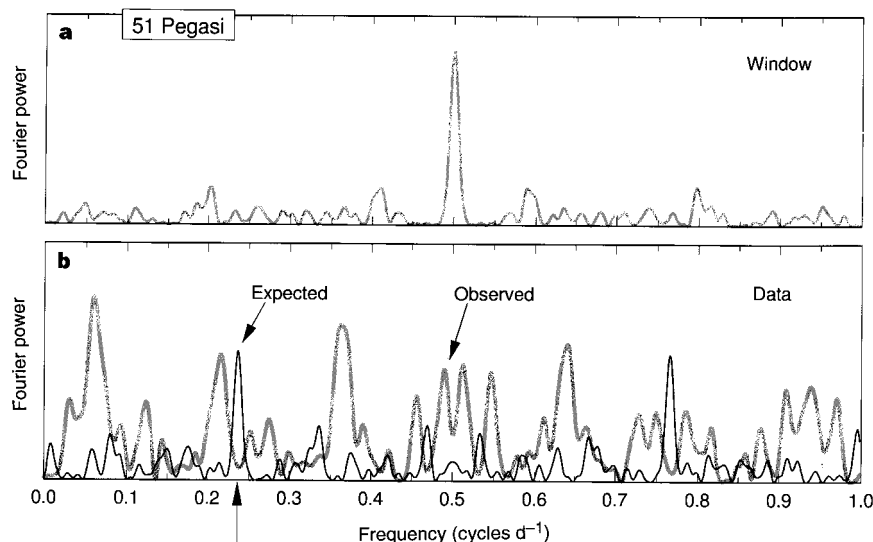


Figure 1 The periodogram of the bisector curvature for the frequency range 0.0–1.0 cycles d⁻¹. **a**, The window pattern for the times of the data spacing. **b**, The thick line labelled 'Observed' shows the periodogram of the data; the thin line labelled 'Expected' is for the expected oscillation with an amplitude of 45 m s⁻¹. No significant power occurs at the frequency of 0.2365 cycles d⁻¹ (indicated by the vertical arrow below the horizontal axis), corresponding to the 4.23-d period seen in the earlier data.

Table 1 Observational data for 51 Pegasi

Date in 1997	JD-2440000*	S/N†	BC‡
20 Jul.	10649.8138	197	74
29 Jul.	10658.7979	349	22
19 Aug.	10679.8262	352	45
29 Aug.	10683.7258	452	86
3 Sep.	10695.7221	228	-102
3 Sep.	10695.8361	320	4
4 Sep.	10696.6444	319	1
4 Sep.	10696.7284	250	11
5 Sep.	10696.8944	420	63
8 Sep.	10700.6382	321	89
14 Sep.	10706.6028	398	30
15 Sep.	10707.6012	378	51
17 Sep.	10709.6333	408	9
17 Sep.	10709.7651	432	-13
18 Sep.	10710.5737	201	-90
24 Sep.	10716.6035	329	63
24 Sep.	10716.7683	452	189
25 Sep.	10717.7084	287	-5
26 Sep.	10718.6261	301	118
27 Sep.	10719.6353	280	83
1 Oct.	10723.6424	221	-45
3 Oct.	10725.6078	478	49
6 Oct.	10728.5462	299	-34
7 Oct.	10729.6049	390	14
8 Oct.	10730.5912	203	4
17 Oct.	10739.6354	267	60
18 Oct.	10740.5826	525	53
25 Oct.	10747.6646	263	129
28 Oct.	10750.5286	294	8
29 Oct.	10751.5951	474	91

* Julian day minus 2440000.

† Signal-to-noise ratios based on exposure level.

‡ Bisector curvature for Fe I 6,253 Å, as defined in the text.

Nor does this result change if lower signal-to-noise exposures are removed. The net result is that I do not see the expected signal.

As Fig. 1 shows noise peaks larger than the expected signal peak, these data alone are not definitive proof that a 4.23-d signal does not exist. However, the same non-detection result was found by Cochran and Hatzes⁵, and by Brown *et al.*^{6,7}, although the relatively low spectral resolving power in the latter set of observations required the development of some innovative analysis techniques.

My 1997 data also yield negative results in similar periodograms of, first, the slope of the line bisector as expressed in the velocity span between $F/F_c = 0.55$ and 0.83 , and second, in the ratio of the depths of the V I line 6,251.83 Å to Fe I 6,252.57 Å that has been used as a temperature index in some previous investigations¹⁰.

As the spectral line variation that I reported previously is not seen in the new observations from the 1997 observing season, there would appear to be only three possible conclusions: (1) the star has stopped pulsating, if indeed non-radial oscillation was the correct interpretation of the variations, or (2) noise is masking the signal in the 1997 data, or (3) the signal is simply not there. The first alternative is ruled out if the radial-velocity variations (that is, simple shifting of spectral line positions), is still observed with the same phase and amplitude. The second alternative cannot be definitively ruled out using my data alone, but when coupled with the observations of Hatzes *et al.*⁵ and Brown *et al.*^{6,7} is hardly tenable. Therefore, it now seems likely that the signal seen in the 1989–96 data was noise, even though the chance of this happening was only one in several hundred (ref. 4). □

Received 12 November; accepted 25 November.

1. Mayor, M. & Queloz, D. A Jupiter-mass companion to a solar-type star. *Nature* **378**, 355–359 (1995).
2. Marcy, G. W. *et al.* The planet around 51 Pegasi. *Astrophys. J.* **481**, 926–935 (1997).
3. Gray, D. F. Absence of a planetary signature in the spectra of the star 51 Pegasi. *Nature* **385**, 795–796 (1997).
4. Gray, D. F. & Hatzes, A. P. Non-radial oscillation in the solar-temperature star 51 Pegasi. *Astrophys. J.* **490**, 412–424 (1997).
5. Hatzes, A. P., Cochran, W. D. & Bakker, E. J. Further evidence for the planet around 51 Pegasi. *Nature* **391**, 154–156 (1998).
6. Brown, T. M. *et al.* A search for line shape and depth variations in 51 Pegasi and τ Boötis. *Astrophys. J.* (submitted).
7. Brown, T. M. *et al.* Exoplanets or dynamic atmospheres? The radial velocity and line shape variations of 51 Pegasi and τ Boötis. *Astrophys. J. Suppl. Ser.* (submitted).

8. Gray, D. F. *Lectures on Spectral-Line Analysis: F, G and K Stars* Ch. 4 (Arva, Ontario, 1988).
9. Gray, D. F. & Desikachary, K. A new approach to periodogram analyses. *Astrophys. J.* **181**, 523–530 (1973).
10. Gray, D. F. Spectral line-depth ratios as temperature indicators in cool stars. *Publ. Astron. Soc. Pacif.* **106**, 1248–1257 (1994).

Acknowledgements. This work was supported by the Natural Sciences and Engineering Research Council of Canada.

Correspondence should be addressed to D.F.G. (e-mail: dfgray@uwo.ca).

Further evidence for the planet around 51 Pegasi

Artie P. Hatzes, William D. Cochran & Eric J. Bakker

McDonald Observatory, The University of Texas at Austin, Austin, Texas 78712, USA

The discovery¹ of a planet around the solar-type star 51 Pegasi marked a watershed in the search for extrasolar planets. Since then, seven other planets have been discovered^{2–6}, of which several have surprisingly short orbital periods, like the planet around 51 Peg. These planets were detected using the indirect technique of measuring variations in the Doppler shifts of lines in the spectra of the primary stars. But it is possible that regular oscillations of the stars themselves (or other effects) could mimic the signature of the planets, particularly the short-period planets. The apparent lack of spectral⁷ and brightness⁸ variations, however, led to widespread acceptance that there is a planet around 51 Peg. This conclusion was challenged by the observation⁹ of systematic variations in the line shapes of 51 Peg, which suggest stellar oscillations¹⁰. If these observations are correct, then there is no need to invoke a planet around 51 Peg to explain the data. Here we report observations of 51 Peg at a much higher spectral resolution than those in ref. 9, in which we find no evidence for systematic changes in the line shapes. The data are most consistent with a planetary companion to 51 Peg.

Observations were made using the coude echelle spectrograph¹¹ of the 2.7-m Harlan J. Smith telescope at the McDonald Observatory. This instrument was used in the configuration that provided a resolving power ($\lambda/\Delta\lambda$) of 220,000, or more than twice that of the data obtained by Gray⁹ and with a wavelength coverage of several hundred (non-contiguous) ångströms. Approximately 120 observations were made over the course of 18 nights in July–September 1997. On each night, 4–10 observations of 51 Peg were made with a typical signal-to-noise ratio of ~ 200 for each spectrum. Information about the spectral line shapes was extracted using the line bisector which is the locus of the midpoints of the stellar absorption line from the line core up to the continuum. The velocity span and the bisector curvature were both measured. The span is the velocity difference between the bisector points taken at two flux levels (in this case 0.48 and 0.85 of the continuum level); the curvature is the difference in the velocity span between the top half (measured between flux points 0.64 and 0.85 of the continuum) and bottom half (measured between flux points 0.42 and 0.64 of the continuum) of the line bisector. (See Gray⁹ and Gray and Hatzes¹⁰ for illustrations of a typical bisector, as well as definitions of the velocity span and curvature.) Both the span and the bisector curvature were examined because it is possible for spectral variations to have large changes in the curvature without significant changes in the velocity span and vice versa. The spectral coverage of our data included the 6,252.57-Å Fe I line used by Gray⁹ as well as four other spectral features of comparable line strength that were suitable for bisector measurements: 5,296.70-Å Cr I, 5,638.27-Å Fe I and 6,141.74-Å Fe I. If the kinematic non-radial pulsation model proposed by Gray and Hatzes¹⁰ is correct, then changes in the lines bisectors for spectral

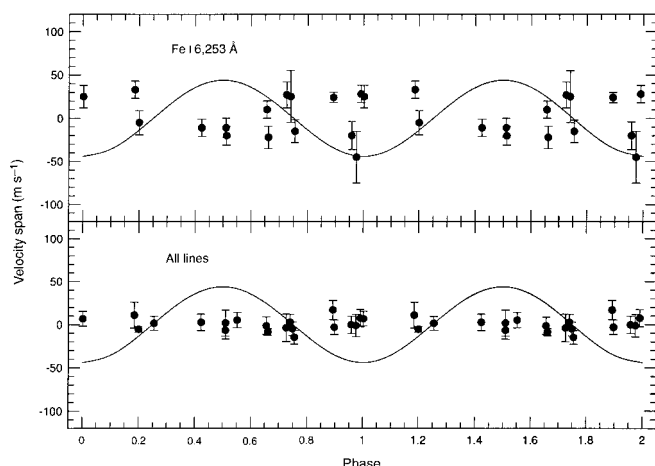


Figure 1 Velocity span measurements of 51 Peg. Top panel, the velocity span of 6,252.57-Å Fe I as a function of phase (data points) according to the ephemeris of Marcy *et al.*¹². The solid line represents the variations reported by Gray⁹. Bottom panel, the average velocity span (data points) for the spectral lines: 5,296.70-Å Cr I, 5,638.27-Å Fe I, 6,141.74-Å Fe I and 6,252.57 Fe I. Again, the solid line represents the predicted curve from Gray⁹.

lines of equal strength should be the same, assuming that the spectral lines are formed at approximately the same depth in the stellar atmosphere. Averaging bisectors from several lines should decrease the noise level in the measurements.

Figure 1 shows the results of our velocity span measurements. The top panel shows the nightly average of the velocity span of the 6,252.57-Å Fe I line, the same feature measured by Gray⁹, as a function of phase. Phases were computed using the orbital ephemeris of Marcy *et al.*¹² (a period of 4.2311 d and an epoch of Julian day 2449797.663) which was also used by Gray and Hatzes¹⁰. The solid line represents the sinusoidal variation predicted by Gray⁹. Each error bar is the standard deviation of individual nightly measurements weighted by the square root of the number of measurements used in the average. The r.m.s. scatter of the measurements about the mean value is $\sim 25 \text{ m s}^{-1}$, which is half of the amplitude of the variations found in 51 Peg by Gray⁹ and is comparable to the r.m.s. scatter for non-variable stars¹⁰. There may be a hint of sinusoidal variability in the measurements, but 180° out-of-phase with the curve predicted by Gray which suggests that this variability is due to noise. This is confirmed in the lower panel of Fig. 1, which shows the average velocity span for all five spectral lines that were examined. The errors shown represent the weighted standard deviation of the span measurements from the five spectral lines. The scatter (standard deviation) of the points shown is $\sim 7 \text{ m s}^{-1}$, well below the amplitude of the Gray prediction (line).

Figure 2 shows the result of our bisector curvature measurements. Once again, the top panel is for the 6,252.57-Å Fe I feature, and the bottom panel is the average for all spectral lines. Errors were computed in the same manner as for Fig. 1. The curve represents the fit to the variations from Gray and Hatzes¹⁰. Spectral variations, above the noise, do not seem to be present in either the measurements for the individual 6,252.57-Å Fe I line or in the average curvature for all five spectral lines. The standard deviation of the averaged measurements (Fig. 2 lower panel) is $\sim 12 \text{ m s}^{-1}$.

The best direct comparison of our data to Gray's is made with the 6,252.57-Å Fe I line, the only spectral feature in common with both data sets. Unfortunately, these alone cannot refute the Gray result with complete certainty. Phasing the data to the 4.2311-d radial velocity period does show a 180° phase shift between the two measurements, and tentatively this would seem to contradict his claim of spectral variability. However, the scatter of our measure-

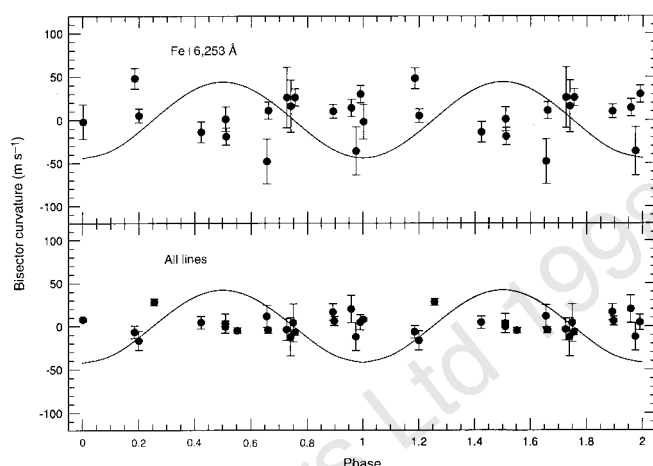


Figure 2 Bisector curve measurements of 51 Peg. Top panel, the bisector curvature of 6,252.57-Å Fe I (data points) as a function of phase. The solid line represents the variations reported by Gray and Hatzes¹⁰. Bottom panel, the average bisector curvature for five spectral lines (data points) along with the predicted curve of Gray and Hatzes¹⁰.

ments are comparable to the amplitude of the variations which may be present. A periodogram analysis (not shown) using our data combined with those of Gray shows this period as the highest peak, and the inclusion of our data reduces the power at secondary peaks somewhat. We note, however, that a periodogram analysis (not shown) of our data only does not show any significant power (false alarm probability $\approx 50\%$) at the appropriate period. A more detailed analysis and its implications will be presented elsewhere.

The averaged bisector measurements, on the other hand, almost certainly refute the result obtained by Gray, but not without caveats. Using the averaged measurements from different lines is valid so long as the line shape variations are the same (and in phase) for all the spectral lines used in the analysis. This is the case for the kinematic pulsation model examined by Gray and Hatzes¹⁰ only if all spectral lines are formed at the same depth in the stellar atmosphere and have similar excitation potentials. However, only the 5,296.70-Å Cr I line has a significantly different excitation potential (0.98 eV) from the other spectral lines (3.6–4.5 eV). If 51 Peg is really a long-period pulsating star, then the vertical (radial) wavelength of the pulsations should be very short, and spectral lines formed at different depths of the stellar atmosphere may show different behaviour with phase in the spectral line shapes. If our line averaging technique is valid, then the spectral variability reported by Gray⁹ is most likely to be an artefact of noise. (Indeed, a Fourier analysis of the data in ref. 9 showed that there was about a 1 in 300 chance that pure noise could produce the observed variability¹⁰.) As the spectral variability of 51 Peg is not confirmed, the non-radial pulsation model for this star proposed by Gray and Hatzes¹⁰ is wrong. We will present elsewhere a more detailed analysis of our work, including line depth ratio measurements, variability of which was also reported by Gray⁹.

Stellar radial velocity measurements provide only an indirect means of finding extrasolar planets and until there is a direct detection (either through imaging or from the spectral signature of the planet) we can never be absolutely certain that 51 Peg indeed has a planet. But in the light of all the observational evidence now available for 51 Peg, including the measurements presented here, the planet hypothesis is the simplest and most plausible explanation for the radial velocity variability reported by Mayor and Queloz¹. □

Received 13 November; accepted 25 November 1997.

1. Mayor, M. & Queloz, D. A Jupiter-mass companion to a solar-type star. *Nature* **378**, 355–359 (1995).
2. Marcy, G. W. & Butler, R. P. A planetary companion to 70 Virginis. *Astrophys. J.* **464**, L147–L151 (1996).

3. Butler, R. P. & Marcy, G. W. A planet orbiting 47 Ursae Majoris. *Astrophys. J.* **464**, L153–L156 (1996).
4. Butler, R. P. *et al.* Three new “51 Peg-type” planets. *Astrophys. J.* **474**, L115–L118 (1997).
5. Cochran, W. D., Hatzes, A. P., Butler, R. P. & Marcy, G. W. The discovery of a planetary companion to 16 Cygni B. *Astrophys. J.* **483**, 5457–5463 (1997).
6. Noyes, R. W. *et al.* A planet orbiting the star ρ Coronae Borealis. *Astrophys. J.* **483**, L111–L114 (1997).
7. Hatzes, A. P., Cochran, W. D. & Johns-Krull, C. M. Testing the planet hypothesis: a search for variability in the spectral line shapes of 51 Peg. *Astrophys. J.* **478**, 374–380 (1997).
8. Henry, G. W., Baliunas, S. L., Donahue, R. A., Soon, W. H. & Saar, S. H. *Astrophys. J.* **474**, 503–510 (1997).
9. Gray, D. F. Absence of a planetary signature in the spectra of the star 51 Pegasi. *Nature* **385**, 795–796 (1997).
10. Gray, D. F. & Hatzes, A. P. Nonradial oscillation in the solar-temperature star 51 Pegasi. *Astrophys. J.* **490**, 412–424 (1997).
11. Tull, R. G., MacQueen, P. J., Sneden, C. & Lambert, D. L. The high-resolution cross-dispersed echelle white pupil spectrometer of the McDonald Observatory 2.7-m telescope. *Publ. Astron. Soc. Pacif.* **107**, 251–264 (1995).
12. Marcy, G. W. *et al.* The planet around 51 Pegasi. *Astrophys. J.* **481**, 926–935 (1997).

Acknowledgements. This work was supported by NASA's Origins of the Solar System Program.

Correspondence should be addressed to A.P.H. (e-mail: artie@astro.as.utexas.edu).

Kondo effect in a single-electron transistor

D. Goldhaber-Gordon^{*†}, Hadas Shtrikman[†], D. Mahalu[†], David Abusch-Magder^{*}, U. Meirav[†] & M. A. Kastner^{*}

^{*} Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

[†] Braun Center for Submicron Research, Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot 76100, Israel

How localized electrons interact with delocalized electrons is a central question to many problems in solid-state physics^{1–3}. The simplest manifestation of this situation is the Kondo effect, which occurs when an impurity atom with an unpaired electron is placed in a metal². At low temperatures a spin singlet state is formed between the unpaired localized electron and delocalized electrons at the Fermi energy. Theories predict^{4–7} that a Kondo singlet should form in a single-electron transistor (SET), which contains a confined ‘droplet’ of electrons coupled by quantum-mechanical tunnelling to the delocalized electrons in the transistor’s leads. If this is so, a SET could provide a means of investigating aspects of the Kondo effect under controlled circumstances that are not accessible in conventional systems: the number of electrons can be changed from odd to even, the difference in energy between the localized state and the Fermi level can be tuned, the coupling to the leads can be adjusted, voltage differences can be applied to reveal non-equilibrium Kondo phenomena⁷, and a single localized state can be studied rather than a statistical distribution. But for SETs fabricated previously, the binding energy of the spin singlet has been too small to observe Kondo phenomena. Ralph and Buhrman⁸ have observed the Kondo singlet at a single accidental impurity in a metal point contact, but with only two electrodes and without control over the structure they were not able to observe all of the features predicted. Here we report measurements on SETs smaller than those made previously, which exhibit all of the predicted aspects of the Kondo effect in such a system.

When the channel of a transistor is made very small and is isolated from its leads by tunnel barriers it behaves in an unusual way. A transistor can be thought of as an electronic switch that is on when it conducts current and off when it does not. Whereas a conventional field-effect transistor, such as one in a computer memory, turns on only once when electrons are added to it, the SET turns on and off again every time a single electron is added to it^{9,10}. This increased functionality may eventually make SETs technologically important.

The unusual behaviour of SETs is a manifestation of the quantization of charge and energy caused by the confinement of the droplet of electrons in the small channel. As similar quantization occurs when electrons are confined in an atom, the small droplet of electrons is often called an artificial atom^{11,12}.

We have fabricated SETs using multiple metallic gates (electrodes) deposited on a GaAs/AlGaAs heterostructure (Fig. 1a) containing a two-dimensional electron gas, or 2DEG. First, the electrons are trapped in a plane by differences in the electronic properties of the heterostructure’s layers. Second, they are excluded from regions of the plane beneath the gates when negative voltages are applied to

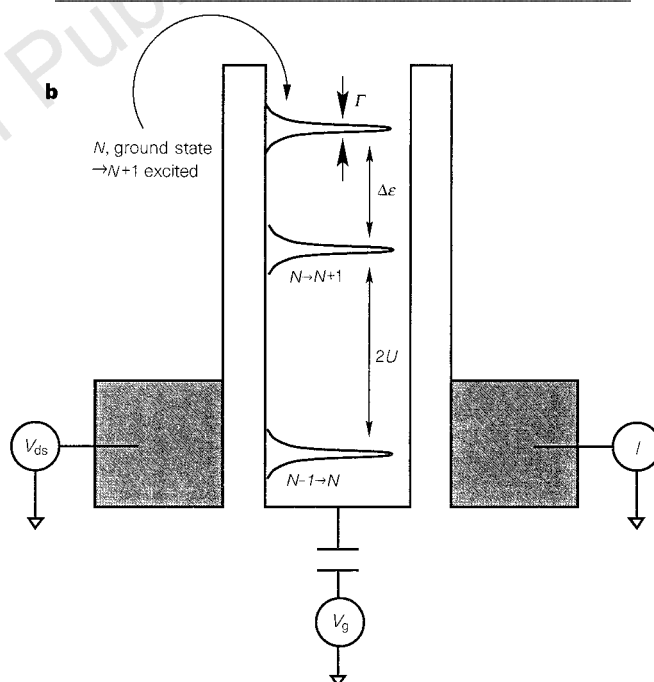
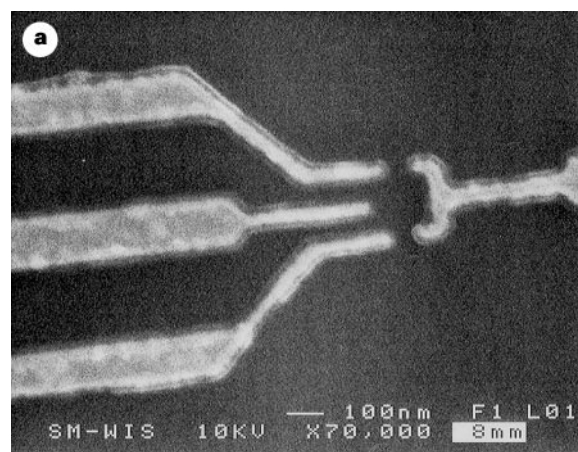


Figure 1 a, Scanning electron microscope image showing top view of sample. Three gate electrodes, the one on the right and the upper and lower ones on the left, control the tunnel barriers between reservoirs of two-dimensional electron gas (at top and bottom) and the droplet of electrons. The middle electrode on the left is used as a gate to change the energy of the droplet relative to the two-dimensional electron gas. Source and drain contacts at the top and bottom are not shown. Although the lithographic dimensions of the confined region are 150 nm square, we estimate lateral depletion reduces the electron droplet to dimensions of 100 nm square. The gate pattern shown was deposited on top of a shallow heterostructure with the following layer sequence grown on top of a thick undoped GaAs buffer: 5 nm $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$, $5 \times 10^{12} \text{ cm}^{-2}$ Si δ -doping, 5 nm $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$, δ -doping, 5 nm $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$, 5 nm GaAs cap (H.S., D.G.-G. and U.M., manuscript in preparation). Immediately before depositing the metal, we etched off the GaAs cap in the areas where the gates would be deposited, to reduce leakage between the gates and the electron gas. **b**, Schematic energy diagram of the artificial atom and its leads. The situation shown corresponds to $V_{ds} < kT/e$, for which the Fermi energies in source and drain are nearly equal, and to a value of V_g near a conductance minimum between a pair of peaks corresponding to the same spatial state. For this case there is an energy cost $\sim U$ to add or remove an electron. To place an extra electron in the lowest excited state costs $\sim U + \Delta\epsilon$.

those gates. This creates a droplet of electrons separated from the leads by tunnel junctions. This basic technique has been used previously^{13–17}. To make our SETs smaller than earlier ones, we have fabricated shallower 2DEG heterostructures (H.S., D.G.-G. and U.M., manuscript in preparation)¹⁸ as well as finer metallic gate patterns by electron-beam lithography. The smaller size of the SETs is critical to our observation of the Kondo effect. (Dimensions are given in Fig. 1a.)

Several important energy scales and their relative sizes determine the behaviour of a SET (see Fig. 1b). At low temperature, the number of electrons N in the droplet is a fixed integer (roughly 50 for our samples). This number may be changed by raising the voltage of a nearby gate electrode which lowers the energy of electrons in the droplet relative to the Fermi level in the leads. The change in energy necessary to add an electron is called U , and in a simple model is the charging energy $e^2/2C$, where C is the capacitance of the droplet. As U is determined by the Coulomb repulsion between pairs of electrons in the droplet, it scales approximately inversely with the droplet's radius.

For small droplets, the quantized energy difference between different spatial electronic states becomes important. We call the typical energy spacing between spatial states $\Delta\epsilon$. Another important energy Γ is the coupling of electronic states on the artificial atom to those on the leads, resulting from tunnelling. When Γ is made greater than $\Delta\epsilon$, the electrons spread from the artificial atom into the leads, and quantization of charge and energy is lost, even at temperature $T = 0$.

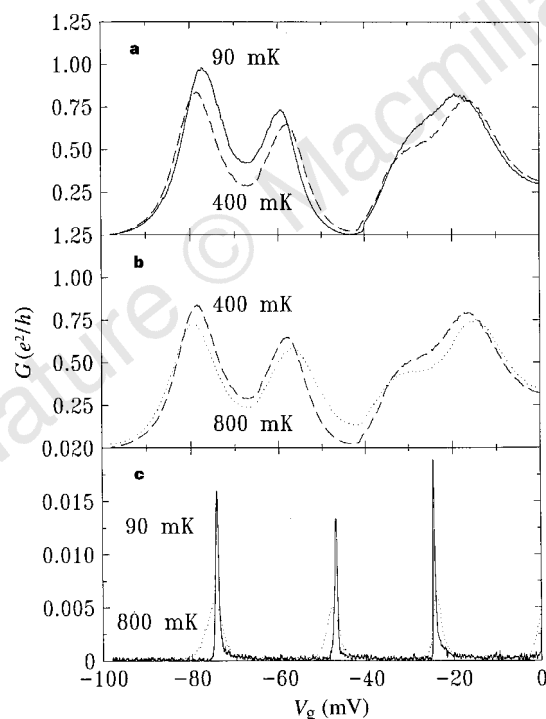


Figure 2 Temperature dependence of zero-bias conductance G through two different spatial states on the droplet. **a**, Paired peaks corresponding to the two spin states for each spatial state become better resolved with increasing temperature from 90 mK (full line) to 400 mK (dashed). The intra-pair valleys become deeper and the peaks become narrower. **b**, From 400 mK (dashed line) to 800 mK (dotted) the paired peaks near $V_g = -70$ mV broaden. The peaks near $V_g = -25$ mV are still becoming better resolved even at 800 mK, as they have larger Γ and hence larger T_K . **c**, When Γ is reduced (as illustrated by shorter and narrower peaks), U increases relative to $\Delta\epsilon$, so peak pairing is no longer evident. Because the Kondo phenomenon is suppressed, peaks become narrower as temperature is decreased at all T down to our base temperature of 90 mK. Full line is for 90 mK, dotted line for 800 mK.

Finally, the energy that determines whether Kondo physics will be visible is kT_K , which is always smaller than Γ . (Here k is Boltzmann's constant and T_K is called the Kondo temperature⁷. By making smaller SETs we have made $\Delta\epsilon$ relatively large, permitting large Γ and thus T_K comparable to accessible temperatures. In semiconductor SETs, Γ can be tuned by changing the voltage on the gates that create the barriers between artificial atom and leads. We find that with our new SETs we can vary Γ slowly as it approaches $\Delta\epsilon$, and thus optimize T_K .

In our experiment, we make two types of measurements. In the first, we apply a voltage of a few microvolts between the two leads to the SET, called the source and drain, and measure the current that flows through the droplet as a function of the voltage V_g on one of the SET's gates (the middle electrode on the left in Fig. 1a). For such low applied voltage ($< kT/e$), current varies linearly with voltage, and the zero-bias conductance can be measured. We use a 10 Hz a.c. excitation, and perform lock-in detection of the current. In the second class of measurements, we add a variable d.c. offset V_{ds} (up to several millivolts) to our a.c. excitation, and again use lock-in detection of current to obtain differential conductance dI/dV_{ds} against V_{ds} .

Varying V_g on a SET typically results in adding an electron to the droplet each time the voltage is increased by a fixed increment proportional to U . As current can flow through the SET only when the occupancy of the island is free to fluctuate between N and $N + 1$, a plot of conductance against V_g shows a series of sharp, periodically spaced peaks^{15–17}. Figure 2c shows this behaviour when Γ is made relatively small. When Γ is large, as in Fig. 2a and b, we find that these peaks form pairs, with large inter-pair spacing and small intra-pair spacing. The two peaks within a pair have comparable widths and heights, whereas between pairs the widths vary significantly. These observations are direct evidence that two electrons of different spin are occupying each spatial state. Between paired peaks, N is odd, whereas between adjacent pairs it is even. As two electrons corresponding to the pair of peaks are added to the same spatial state, the intra-pair spacing is determined by U . However, when N is even (between pairs) the next electron must be placed in a different spatial state, so the interpair spacing is determined by $U + \Delta\epsilon$.

When $\Gamma \leq kT$ the peaks in conductance against gate voltage become narrower and larger with decreasing T , as illustrated in Fig. 2c, saturating at a width and height determined by Γ . This behaviour results from the sharpening of the Fermi distribution^{15,18}. Figure 2 illustrates that this pattern is followed for large Γ , as well, on the outside edges of paired peaks at all T . However, inside pairs the peaks become narrower as T is reduced from 800 mK to 400 mK, but then broaden again at low temperatures, resulting in increased conductance in the intra-pair valley. Thus, not only the peak spacing but also the mechanism of conduction itself is different within pairs from between pairs. The enhancement of linear conductance at low temperature for odd but not even N is a manifestation of Kondo physics¹⁹. If N is odd there is an unpaired electron with a free spin which can form a singlet with electrons at the Fermi level in the leads. This coupling results in a greater density of states at the Fermi level of the leads, and hence higher conductance^{4,5}. Raising the temperature destroys the singlet and attenuates the conductance.

Another aspect of the Kondo effect is the sensitivity of the excess conductance between the pair of peaks to the difference in Fermi levels in the two leads. The extra electron in the droplet couples to electrons in both leads, giving a greater density of states at both Fermi levels^{7,20,21}. When the applied voltage is large, separating the Fermi levels in the two leads, the electrons at the Fermi level in the higher energy lead can no longer resonantly tunnel into the enhanced density of states in the lower energy lead, so the extra conductance is suppressed. This can be seen in Fig. 3, a plot of differential conductance against V_{ds} . The enhanced conductance is suppressed by a bias of ~ 0.1 mV in either direction.

Figure 3 illustrates how this nonlinear conductance measurement

also offers a complementary way of seeing the suppression of the zero-bias enhancement with temperature. By 600 mK, the Kondo resonance has almost completely disappeared between the peak pair near $V_g = -70$ mV.

A magnetic field also alters the Kondo effect. Applying a magnetic field splits the unpaired localized electron state into a Zeeman doublet separated by the energy $g\mu_B B$. This also splits the enhanced density of states at the Fermi level into two peaks with energies $g\mu_B B$ above and below the Fermi level⁷. When the Fermi level of one lead is raised or lowered by a voltage $g\mu_B B/e$ relative to the other lead, electrons can tunnel into the Kondo-enhanced density of states. In differential conductance at high magnetic field, we thus see peaks at $\pm g\mu_B B/e$ (see Fig. 3). The splitting of peaks in differential conductance by *twice* $g\mu_B B$ (compared to the spin splitting of the localized state, $g\mu_B B$) provides a distinctive signature of Kondo physics. Because the peaks are broad and overlapping, measuring the distance between their maxima may underestimate their splitting at lower magnetic fields. At 7.5 T, when the peaks no longer overlap, we find their splitting to be 0.033 ± 0.002 meV T^{-1} . In comparison with measurements on bulk GaAs, this is significantly smaller than $2g\mu_B B$ yet larger than $g\mu_B B = 0.025$ meV T^{-1} . Electron spin resonance measurements have found²² that spin splittings in a two-dimensional electron gas are suppressed compared with values for bulk samples, sometimes by as much as 35%. Thus, our measurement is consistent with a splitting of $2g\mu_B B$.

Figure 4 shows the differential conductance at low temperature and zero magnetic field as a function of both V_{ds} and V_g . This range of V_g spans the better-resolved pair of peaks near $V_g = -70$ mV in Fig. 2 as well as the valleys on either side of it. The bright diagonal lines result from strong peaks in dI/dV_{ds} (outside the range of V_{ds} shown in Fig. 3) marking the values of V_{ds} and V_g where N can change to $N+1$ or $N-1$. The slopes of these lines contain information about the relative capacitances of gates and leads to the droplet of electrons. Together with the intra-pair spacing of peaks in zero-bias conductance as a function of V_g at 800 mK (Fig. 2), these relative capacitances give a value for U of about

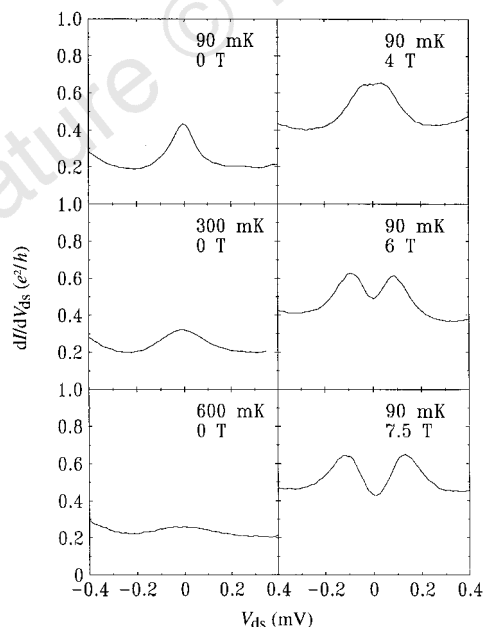


Figure 3 Temperature and magnetic field dependence of the zero-bias Kondo resonance measured in differential conductance. Increasing temperature suppresses the resonance, whereas increasing magnetic field causes it to split into a pair of resonances at finite bias. These scans were made with gate voltage halfway between the paired peaks near $V_g = -70$ mV in Fig. 2. Because the energy of the spatial state we are studying changes with magnetic field, the valley between our two peaks occurs at a different gate voltage for each value of magnetic field.

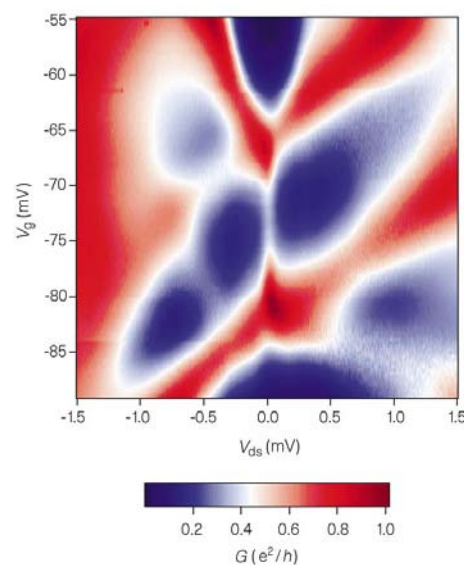


Figure 4 Differential conductance on a colour scale as a function of both V_g and V_{ds} . The white vertical line between the two maxima indicates that there is a zero-bias peak for odd N only.

0.6 meV. The ratio $\Gamma/\Delta\epsilon$ is of order unity, so electron wavefunctions in the droplet extend somewhat into the leads, suppressing U . When Γ is reduced for this same device (as in Fig. 2c), U increases to ~ 2 meV. From the difference between inter- and intra-pair spacing in Fig. 2 we find $\Delta\epsilon \approx U$. To determine Γ , we set V_g so that the spatial state corresponding to the paired peaks is empty. Then the width of the peak in a plot of dI/dV_{ds} against V_{ds} corresponding to tunnelling into the empty state is the bare level width $\Gamma \approx 0.2$ meV. This is unmodified by Kondo-enhanced tunnelling, which can only occur when the state is singly occupied in equilibrium. All three quantities U , $\Delta\epsilon$ and Γ are considerably larger than $kT = 0.0078$ meV at base temperature.

A white vertical line at $V_{ds} = 0$ in Fig. 4 shows that the zero-bias conductance is enhanced everywhere between the paired peaks but not outside the pair. We find, in fact, that there is a zero-bias suppression just outside the pair, especially to the side towards negative gate voltage where our Kondo level is unoccupied²⁰. The sharpness of the Kondo peak compared with other features is a dramatic illustration that we have, indeed, observed the Kondo effect in our SET. □

Received 29 July; accepted 14 October 1997.

- Lawrence, J. M. & Mills, D. L. Recent progress in heavy fermion/valence fluctuation physics: Introduction. *Comments Cond. Mat. Phys.* **15**, 163–174 (1991).
- Grüner, G. & Zawadowski, A. Magnetic impurities in non-magnetic metals. *Rep. Prog. Phys.* **37**, 1497–1583 (1974).
- Varma, C. M. Mixed-valence compounds. *Rev. Mod. Phys.* **48**, 219–238 (1976).
- Ng, T. K. & Lee, P. A. On-site Coulomb repulsion and resonant tunneling. *Phys. Rev. Lett.* **61**, 1768–1771 (1988).
- Glazman, L. I. & Raikh, M. E. Resonant Kondo transparency of a barrier with quasilocal impurity states. *JETP Lett.* **47**, 452–455 (1988).
- Meir, Y., Wingreen, N. S. & Lee, P. A. Low-temperature transport through a quantum dot: the anderson model out of equilibrium. *Phys. Rev. Lett.* **70**, 2601–2604 (1993).
- Wingreen, N. S. & Meir, Y. Anderson model out of equilibrium: Noncrossing approximation approach to transport through a quantum dot. *Phys. Rev. B* **49**, 11040–11052 (1994).
- Ralph, D. C. & Buhrman, R. A. Kondo-assisted and resonant tunneling via a single charge trap: a realization of the Anderson model out of equilibrium. *Phys. Rev. Lett.* **72**, 3401–3404 (1994).
- Fulton, T. A. & Dolan, G. J. Observation of single-electron charging effects in small tunnel junctions. *Phys. Rev. Lett.* **59**, 109–112 (1987).
- Meirav, U. & Foxman, E. B. Single-electron phenomena in semiconductors. *Semiconductor Sci. Technol.* **10**, 255–284 (1995).
- Kastner, M. Artificial atoms. *Phys. Today* **46**, 24–31 (1993).
- Ashoori, R. C. Electrons in artificial atoms. *Nature* **379**, 413–419 (1996).
- van Wees, B. J. et al. Quantized conductance of point contacts in a two-dimensional electron gas. *Phys. Rev. Lett.* **60**, 848–851 (1988).
- Wharam, D. A. et al. One-dimensional transport and the quantisation of the ballistic resistance. *J. Phys. C: Solid State Phys.* **21**, L209–214 (1988).
- Meirav, U., Kastner, M. A. & Wind, S. J. Single-electron charging and periodic conductance resonances in GaAs nanostructures. *Phys. Rev. Lett.* **65**, 771–774 (1990).

16. Johnson, A. T. *et al.* Zero-dimensional states and single-electron charging in quantum dots. *Phys. Rev. Lett.* **69**, 1592–1595 (1992).
17. Weis, J., Haug, R. J., Klitzing, K. v. & Ploog, K. Competing channels in single-electron tunneling through a quantum dot. *Phys. Rev. Lett.* **71**, 4019–4022 (1993).
18. Foxman, E. B. *et al.* Crossover from single-level to multilevel transport in artificial atoms. *Phys. Rev. B* **50**, 14193–14199 (1994).
19. Wan, Y., Phillips, P. & Li, Q. Suppression of the Kondo effect in quantum dots by even–odd asymmetry. *Phys. Rev. B* **51**, R14782–14785 (1995).
20. König, J., Schmid, J., Schoeller, H. & Schön, G. Resonant tunneling through ultrasmall quantum dots: Zero-bias anomalies, magnetic-field dependence, and boson-assisted transport. *Phys. Rev. B* **54**, 16820–16837 (1996).
21. Hershfield, S., Davies, J. H. & Wilkins, J. W. Probing the Kondo resonance by resonant tunneling through an Anderson impurity. *Phys. Rev. Lett.* **67**, 3720–3723 (1991).
22. Dobers, M., von Klitzing, K. & Weimann, G. Electron-spin resonance in the two-dimensional electron gas of GaAs–Al_xGa_{1–x}As heterostructures. *Phys. Rev. Lett.* **38**, 5453–5456 (1988).

Acknowledgements. We thank G. Bunin for help with fabrication, N. Y. Morgan for help with measurements, and I. Aleiner, R. C. Ashoori, M. H. Devoret, D. Esteve, D. C. Glatli, A. S. Goldhaber, L. Levitov, K. Matveev, N. Zhitenev, and especially N. S. Wingreen and Y. Meir for discussions. This work was supported by the US Joint Services Electronics Program under contract from the Department of the Army, Army Research Office. D.G.-G. thanks the students and staff of the Weizmann Institute's Braun Center for Submicron Research for their hospitality during his stay there, and the Hertz Foundation for fellowship support. D.A.-M. thanks the Lucent Technologies Foundation for fellowship support.

Correspondence should be addressed to M.A.K. (e-mail: mkastner@mit.edu).

Metal–insulator transition induced by oxygen isotope exchange in the magnetoresistive perovskite manganites

N. A. Babushkina*, L. M. Belova*, O. Yu. Gorbenko†, A. R. Kaul†, A. A. Bosak†, V. I. Ozhogin* & K. I. Kugel‡

* RRC Kurchatov Institute, 123182 Moscow, Russia

† Chemistry Department, Moscow State University, 119899 Moscow, Russia

‡ Scientific Center for Applied Problems in Electrodynamics, 127412 Moscow, Russia

Perovskite manganites derived from LaMnO₃ have recently become the subject of intensive study following the discovery of ‘colossal’ magnetoresistance (a magnetically induced change in electrical resistance of up to several orders of magnitude) in several members of this family of compounds¹. The manganites exhibit a broad range of electronic and magnetic phases, ranging from low-resistance ferromagnetic metals to high-resistance insulators, which are extremely sensitive to variation of composition², temperature and pressure³. A recent study showed that such sensitivity also extends to oxygen isotope exchange⁴: replacing ¹⁶O with ¹⁸O in La_{0.8}Ca_{0.2}MnO₃ produces an unusually large change in the magnetic properties (a 21-kelvin decrease in the Curie temperature). The magnitude of this isotope shift is evidence for the essential role played by electron–phonon coupling⁵ in determining the transport properties of these materials. Here we show that this sensitivity to oxygen isotope exchange can be even more extreme. In its normal state, the compound La_{0.175}Pr_{0.525}Ca_{0.3}MnO₃ undergoes an insulator-to-metal transition as it is cooled below ~95 K. But we find that, after substituting ¹⁸O for ¹⁶O, the compound remains an insulator down to 4.2 K, so providing a vivid demonstration of the importance of lattice vibrations in these materials.

Here we consider the perovskite manganites R_{1–x}M_xMnO₃ (where R³⁺ is a rare earth cation, M is a doubly charged cation of large ionic radius, and both R and M fill A positions of ABO₃ perovskite lattice; ref. 1). In this system, colossal magnetoresistance (CMR) occurs in the temperature range of the phase transitions resulting in the formation of a ferromagnetic phase: only the ferromagnetic phase shows the metal-like conductivity. The most likely explanation of such behaviour relates the electrical conductivity to the strong intra-atomic spin coupling of *t*_{2g} and *e*_g electrons, and double

exchange between neighbouring Mn³⁺ and Mn⁴⁺ ions assisted by 2*p* orbitals of O. At high temperatures, all R_{1–x}M_xMnO₃ phases are paramagnetic and semiconducting. With a decrease of temperature, a variety of phase transformations occurs depending on the doping level (*x*) and average ionic radius of the cations in the A position (*r*_A). In the range *x* = 0.2–0.5 the ferromagnetic transition takes place if the tolerance factor *t* (where *t* = *d*_{A–O}/√2*d*_{Mn–O}, and *d* denotes a interatomic distance) exceeds the critical value 0.91 as with La_{1–x}Ca_xMnO₃ (ref. 2); this value has been found empirically using Shannon ionic radii. The Mn–O bond length varies insignificantly with changes of R and M at the constant Mn⁴⁺/Mn³⁺ ratio, whereas the A–O bond length is a function of *r*_A. MnO₆ octahedra forming a three-dimensional array in the perovskite structure can be considered nearly invariant. These octahedra tilt and rotate to follow the reduction of space around the A cation if *t* decreases². This buckling results in a decrease of the Mn–O–Mn bond angle *θ*, which disrupts the overlapping of 2*p* orbitals of O and *e*_g orbitals of Mn ions. If the overlapping is insufficient, then double exchange is no longer possible in Mn–O–Mn chains. The localized holes at *e*_g orbitals tend to lead to the formation of a charge-ordered (CO) state of the type found in La_{0.5}Ca_{0.5}MnO₃ (ref. 6); the material also becomes antiferromagnetic. Such a situation occurs in Pr_{1–x}Ca_xMnO₃ (*t* < 0.91 for *x* = 0.2–0.4), which has been intensively studied^{7,8}.

The behaviour mentioned above suggests that, besides double exchange, the electron–phonon coupling must play an essential role in the CMR manganites. For instance, the stretching vibration of the Mn–O–Mn chain would change angle *θ* (which controls the double exchange) and consequently the electrical conductivity of the chain. For *t* ≈ 0.91, one can imagine a material which still has metal-like conductivity owing to the stretching vibration, where the angle *θ* periodically reaches a sufficient value for effective double exchange. If one suppresses the vibration, then the material under study is expected to become a CO insulator. We therefore considered the possibility of producing a pronounced variation of the electrical conductivity by substitution of ¹⁸O for ¹⁶O in R_{1–x}M_xMnO₃ at *t* ≈ 0.91 owing to the strong electron–phonon coupling. For the first time, a 21 K decrease of the Curie temperature *T*_C after isotope exchange ¹⁸O → ¹⁶O has been reported by Zhao *et al.*⁴ for La_{0.8}Ca_{0.2}MnO₃ with *t* ≈ 0.92. For compositions with higher *t* the isotope effect was less profound; in the case of La_{0.945}Mn_{0.945}O₃, a *T*_C shift of ~12 K was observed⁹. These results suggest that a stronger oxygen isotope effect could be obtained by decreasing *t*.

The system (La_{1–y}Pr_y)_{0.7}Ca_{0.3}MnO₃ should be suitable for testing this approach to the production of a strong isotope effect. On the one hand, in this series of solid solutions *t* passes through critical value 0.91. Starting at ~260 K for La_{0.7}Ca_{0.3}MnO₃, both *T*_C and maximum resistivity temperature (*T*_p) decrease rapidly with increasing Pr content: Pr_{0.7}Ca_{0.3}MnO₃ is a CO insulator. On the other hand, (La_{1–y}Pr_y)_{0.7}Ca_{0.3}MnO₃ is a convenient system to work with because the oxygen stoichiometry is rather insensitive to variations of the heat treatment (temperature and partial pressure of O₂)². So the Mn⁴⁺/Mn³⁺ ratio is constant in the series. The isotope effect observed by us for the composition La_{0.175}Pr_{0.525}Ca_{0.3}MnO₃ has been far in excess of our expectations, as not merely a shift of *T*_p but also a metal–insulator transition due to the isotope exchange was established.

To prepare ceramic samples, aqueous solutions of La(NO₃)₃, Pr(NO₃)₃, Ca(NO₃)₂ and Mn(NO₃)₂ were mixed together in the proper ratio, and ash-free paper filters were soaked with the solution. The paper was dried at 120 °C and then burned: the product was annealed at 700 °C for 2 h. Pellets pressed from the powder residue were sintered at 1,200 °C for 12 h in air. Ceramics of composition La_{0.175}Pr_{0.525}Ca_{0.3}MnO₃ were obtained, and were found to be single phase according to X-ray diffraction with orthorhombic lattice parameters *a* = 0.5436(2) nm, *b* = 0.5461(2) nm, *c* = 0.7686(3) nm at 300 K.

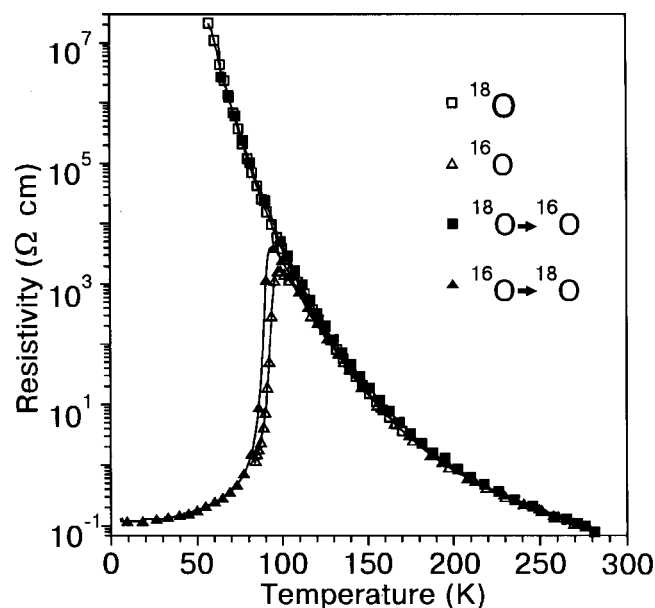


Figure 1 Temperature dependence of resistivity for $\text{La}_{0.175}\text{Pr}_{0.525}\text{Ca}_{0.3}\text{MnO}_3$ ceramic samples after the first isotope exchange (^{18}O and ^{16}O), and after the isotope back-exchange ($^{18}\text{O} \rightarrow ^{16}\text{O}$ and $^{16}\text{O} \rightarrow ^{18}\text{O}$; here $\rightarrow$ means 'replaces'). The reversibility of the metal-insulator transition induced by the isotope exchange is clear. For clarity, only each tenth experimental point is shown.

Two $1 \times 1 \times 8 \text{ mm}^2$ bars were cut from the sintered pellet of $\text{La}_{0.175}\text{Pr}_{0.525}\text{Ca}_{0.3}\text{MnO}_3$ and put into alumina boats which were placed in two quartz tubes mounted in the furnace. The quartz tubes formed parts of two identical closed loops where enforced circulation of gas was applied¹⁰. Both samples were treated simultaneously: one sample was heated in an atmosphere of $^{16}\text{O}_2$, the other was heated in $^{18}\text{O}_2$ (oxygen with molar fraction of $^{18}\text{O}_2$ 85%). The diffusion annealing was carried out for 48 h at 950 °C under an oxygen pressure of 1 bar. The ^{18}O concentration in the samples was determined by measurement of the weight change of the samples after isotope enrichment, and completion of the isotope exchange was established. The composition was adjusted to $\text{La}_{0.175}\text{Pr}_{0.525}\text{Ca}_{0.3}\text{Mn}(\text{}^{18}\text{O}_{0.85}\text{}^{16}\text{O}_{0.15})_3$.

Electrical resistivity was measured with a standard four-point configuration down to 4.2 K; the highest resistance which could be measured with the experimental set-up was 1 GΩ. The resistivity measurements in a magnetic field applied along the bar sample were accomplished only at the cooling stage. To avoid hysteresis^{7,11}, the samples were heated above 180 K before a variation of the magnetic field.

The resistivity curves for samples treated in $^{16}\text{O}_2$ and $^{18}\text{O}_2$ differed drastically (Fig. 1). The former showed a maximum of resistivity at 95 K (typical for a ferromagnetic phase transition in $\text{R}_{1-x}\text{M}_x\text{MnO}_3$); the latter were insulating down to 4.2 K.

To make sure that the effect was due to isotope exchange, we carried out the isotope back-exchange; completion of the exchange was proved again by change of the sample weight. The sample which had been first saturated by ^{18}O and was insulating thereafter, became metallic below 95 K after subsequent annealing in $^{16}\text{O}_2$. The sample which had been first treated in $^{16}\text{O}_2$ and was metallic below 95 K thereafter, became an insulator after the subsequent treatment in $^{18}\text{O}_2$ (Fig. 1).

The application of a magnetic field is known to provide 'melting' of the CO state in $(\text{La}_{1-y}\text{Pr}_y)_{0.7}\text{Ca}_{0.3}\text{MnO}_3$, resulting in a metamagnetic phase transition with a drastic decrease of the electrical resistivity^{7,11}. We evaluated the stability of the CO state (induced by the isotope exchange) to an applied magnetic field. We found that a magnetic field of 1 T was not enough to suppress the CO state in $\text{La}_{0.175}\text{Pr}_{0.525}\text{Ca}_{0.3}\text{Mn}(\text{}^{18}\text{O}_{0.85}\text{}^{16}\text{O}_{0.15})_3$ in spite of the more than

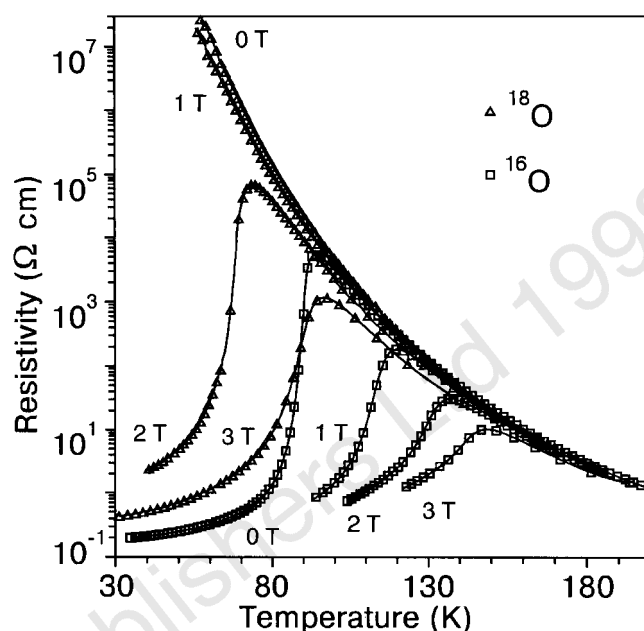


Figure 2 Temperature dependence of resistivity for $\text{La}_{0.175}\text{Pr}_{0.525}\text{Ca}_{0.3}\text{MnO}_3$ under various magnetic fields. The magnetic field of 2 T suppresses the CO state of the ^{18}O sample, resulting in the metal-insulator transition.

30 K upward shift of T_p in the case of $\text{La}_{0.175}\text{Pr}_{0.525}\text{Ca}_{0.3}\text{Mn}(\text{}^{16}\text{O})_3$. Nevertheless, in a field of 2 T the 'melting' of the CO state was clearly detected (Fig. 2). T_p increased significantly with the increase of the applied magnetic field for both samples. The isotope shift of T_p was 63 K in the field of 2 T, and 54 K in the field of 3 T, which is considerably more than isotope effects reported for $\text{R}_{1-x}\text{M}_x\text{MnO}_3$. The resistivity curves at higher temperature are practically independent of the isotope exchange (Fig. 2).

We now consider the nature of the electron-phonon coupling responsible for the huge isotope effect. In general, there are three kinds of distortions in the perovskite structure under study^{7,12}: (1) buckling of MnO_6 octahedra as discussed above, which determines orthorhombic symmetry of the lattice as a function of t ; (2) Jahn-Teller distortion which has an influence on the ratio of the lattice parameters (in particular, $c/\sqrt{2} < a < b$ for the structure under study)¹³; and (3) charge ordering giving rise to the superlattice in the [010] direction. The material under study is very close to being in the charge-ordered state. The dynamic lattice distortions due to the Jahn-Teller effect, and buckling of MnO_6 octahedra accompanied by significant electron-phonon coupling, may be responsible for the huge isotope effect. Resistivity data at higher temperatures imply that the isotope effect is connected more with the buckling in Mn-O-Mn chains than with distortion of the MnO_6 octahedra, as in the latter case the bandgap should also be influenced by the isotope exchange. In any case, the isotope effect should be related to such lattice distortions which are characterized by strong elastic anharmonicity. □

Received 18 July; accepted 11 November 1997.

1. Rao, C. N. R., Cheetham, A. K. & Mahesh, R. Giant magnetoresistance and related properties of rare earth manganites and other oxide systems. *Chem. Mater.* **8**, 2421-2432 (1996).
2. Hwang, H. Y., Cheong, S. W., Radaelli, P. G., Marezio, M. & Batlogg, B. Lattice effects on the magnetoresistance in doped LaMnO_3 . *Phys. Rev. Lett.* **75**, 914-917 (1995).
3. Zhou, J.-S., Archibald, W. & Goodenough, J. B. Identification of a new type of electronic state in the magnetoresistive orthomanganites. *Nature* **381**, 770-772 (1996).
4. Zhao, G. M., Conder, K., Keller, H. & Mueller, K. A. Giant oxygen isotope shift in the magnetoresistive perovskite $\text{La}_{1-x}\text{Ca}_x\text{MnO}_{3+y}$. *Nature* **381**, 676-678 (1996).
5. Millis, A. J., Littlewood, P. B. & Shraiman, B. I. Double exchange alone does not explain the resistivity of $\text{La}_{1-x}\text{Sr}_x\text{MnO}_3$. *Phys. Rev. Lett.* **74**, 5144-5147 (1995).
6. Chen, C. H. & Cheong, S.-W. Commensurate to incommensurate charge ordering and its real-space images in $\text{La}_{0.5}\text{Ca}_{0.5}\text{MnO}_3$. *Phys. Rev. Lett.* **76**, 4042-4045 (1996).

7. Yoshizawa, H., Kawano, H., Tomioka, Y. & Tokura, Y. Magnetic field-induced metal-insulator transition in $\text{Pr}_{0.7}\text{Ca}_{0.3}\text{MnO}_3$. *J. Phys. Soc. Jpn* **65**, 1043–1052 (1996).
8. Lees, M. R., Barratt, J., Balakrishnan, G. & McPaul, D. Influence of charge and magnetic ordering on the insulator-metal transition in $\text{Pr}_{1-x}\text{Ca}_x\text{MnO}_3$. *Phys. Rev. B* **52**, 14303–14307 (1995).
9. Zhao, G. M., Hunt, M. B. & Keller, H. Strong oxygen-mass dependence of the thermal expansion coefficient in the manganites $(\text{La}_{1-x}\text{Ca}_x)_{1-y}\text{Mn}_{1-y}\text{O}_3$. *Phys. Rev. Lett.* **78**, 955–958 (1997).
10. Babushkina, N. A. *et al.* Anomalous isotope effect in high-temperature superconductor LSCO with 3d impurities. *Physica C* **272**, 257–268 (1996).
11. Gorbenko, O. Yu., Kaul, A. R., Babushkina, N. A. & Belova, L. M. Giant magnetoresistive thin films of $(\text{La,Pr})_{0.7}(\text{Ca,Sr})_{0.3}\text{MnO}_3$ prepared by aerosol MOCVD. *J. Mater. Chem.* **7**, 747–752 (1997).
12. Jirak, Z., Krupicka, S., Šimša, Z., Dlouha, M. & Vratilav, S. Neutron diffraction study of $\text{Pr}_{1-x}\text{Ca}_x\text{MnO}_3$. *J. Magn. Magn. Mater.* **53**, 153–166 (1985).
13. Garcia-Muñoz, J. L., Saadi, M., Fontcuberta, J. & Rodriguez-Carvajal, J. Reduction of the Jahn–Teller distortion at the insulator-to-metal transition in mixed valence manganites. *Phys. Rev. B* **55**, 34–37 (1997).

Acknowledgements. This work was supported in part by the Russian Foundation for Basic Researches.

Correspondence and requests for materials should be addressed to O.Yu.G. (e-mail: gorbenko@htsc.chem.msu.su).

Controlling polymer shape through the self-assembly of dendritic side-groups

V. Percec*, C.-H. Ahn*, G. Ungar†, D. J. P. Yeardley†, M. Möller‡ & S. S. Sheiko‡

* The W. M. Keck Laboratories for Organic Synthesis, Department of Macromolecular Science, Case Western Reserve University, Cleveland, Ohio 44106-7202, USA

† Department of Engineering Materials and Centre for Molecular Materials, University of Sheffield, Sheffield S1 3JD, UK

‡ Organische Chemie II/Makromolekulare Chemie, Universität Ulm, D-89069 Ulm, Germany

The chain conformation of polymers plays an important role in controlling their phase behaviour and associated material properties. In the case of flexible polymers, conformation is controlled by the degree of polymerization (DP), with low-DP polymers having extended polymer chains and high-DP polymers adopting random-coil conformations in solution and the bulk amorphous state¹, and folded conformations in the crystalline phase². Exceptions to this general rule are polymers that contain structurally rigid building blocks, or that are subjected to directional shear forces during solidification. The backbones of semi-flexible and rigid rod-like polymers, for example, are always extended in liquid crystalline and crystalline phases^{3–5}, and gel-spun flexible polymers form extended-chain crystals⁶. Here we report a general strategy for the rational control of polymer conformation through the self-assembly of quasi-equivalent monodendritic (branched) side-groups attached to flexible backbones. At low DPs, the conical monodendrons assemble to produce a spherical polymer with random-coil backbone conformation. On increasing the DP, the self-assembly pattern of the monodendritic units changes to give cylindrical polymers with extended backbones. This correlation between polymer conformation and DP is opposite to that seen in most synthetic and natural macromolecules. We anticipate that our strategy will provide new approaches for the rational design of organized supramolecular materials^{6–9} in areas such as nanotechnology, functional films and fibres, molecular devices, and membranes, expanding the synthetic and technological uses of dendritic building blocks^{7,10–15}.

Quasi-equivalent building blocks are chemically identical sub-units which self-control their shape by switching between different conformational states during the process of self-assembly¹⁶. Therefore, quasi-equivalence discards the constraints of strict equivalence while retaining the physical essentials of specificity to allow the assembly of chemically identical units into a system of minimum free energy. Classic examples of quasi-equivalent units are the

protein coats which surround a nucleic acid in the self-assembly of icosahedral viruses^{16–19} (Fig. 1). Icosahedral viruses, best approximated by a spherical shape, can have a coil conformation of their own nucleic acid whereas the nucleic acid penetrating through the centre of a rod-like virus has a helical conformation. Figure 1 outlines the principles we advance for the design of cylindrical and spherical macromolecular systems by using monodendrons as synthetic analogues of quasi-equivalent proteins. We have elaborated a convergent¹⁰ synthesis for flat-tapered first-generation monodendrons¹¹ which self-assemble⁶ into supermolecular cylinders¹¹, and of conical and hemispherical monodendrons which self-assemble into spherical supermolecular dendrimers¹². The rationale used in the design of the monodendrons themselves (that is, monodendrons that are not attached to a polymer backbone), their self-assembly into cylindrical and spherical supermolecular dendrimers and subsequent self-organization into hexagonal columnar (Φ_h)¹¹ and cubic (Cub)¹² thermotropic liquid crystal (LC) phases (Fig. 1) were detailed previously. Their analysis by X-ray diffraction (XRD)^{11,12} and transmission electron microscopy (TEM)²⁰ allowed the determination of the shape, size, internal structure and the direct visualization of both spherical and cylindrical supermolecules. The attachment of flat-tapered monodendrons to a flexible polymer backbone induces a helical

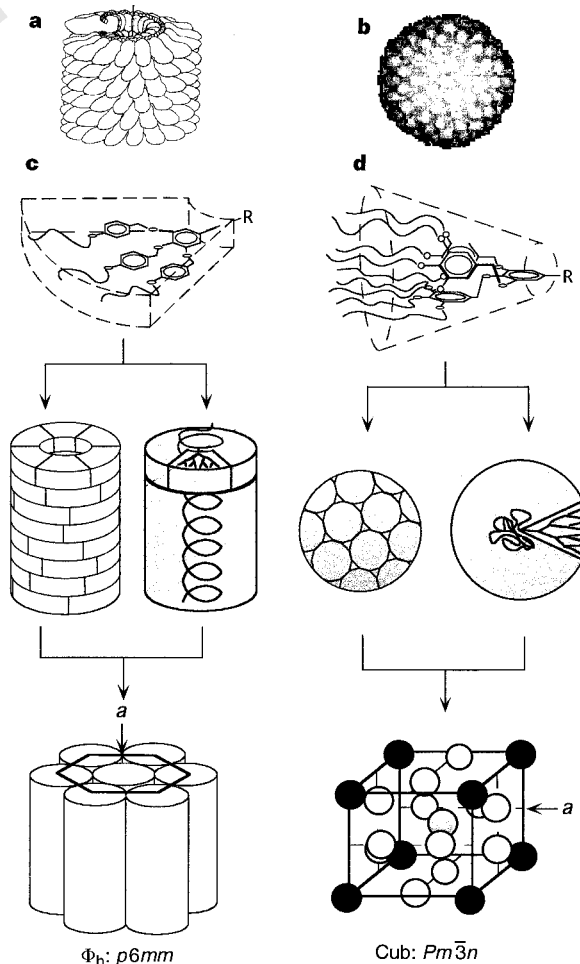


Figure 1 Natural and synthetic supramolecular systems with cylindrical and spherical shapes: **a**, tobacco mosaic virus; **b**, icosahedral virus; synthetic analogue of **a** and **b** have been self-assembled from tapered (**c**) and conical (**d**) monodendrons containing endoreceptors or polymer backbones as R groups and their subsequent self-organization in $p6mm$ Φ_h and $Pm\bar{3}n$ cubic phases, respectively.

conformation of the backbone by the self-assembly of its tapered-side groups into a cylindrical structure²¹.

Polymers with low and high DPs were synthesized by the radical polymerization of the conical monodendron 12G2-AG¹² functionalized with, respectively, styrene (12G2-AG-S) and methacrylate (12G2-AG-MA) groups (Fig. 2 and Table 1). Their mechanism of polymerization was reported elsewhere²⁷. Differential scanning calorimetry (DSC) traces of the low-DP 12G2-AG-PS and 12G2-AG-PMA are shown in Fig. 3. Both polymers show melting (at -14 and -11 °C, respectively), crystallization (at -19 and -16 °C) and glass transition (at 34 and 35 °C). An exothermic peak (at 59 and 70 °C, respectively) associated with the formation of an optically isotropic cubic phase of $Pm\bar{3}n$ symmetry (demonstrated by XRD, Table 1), is observed on heating. This is followed by isotropization (at 79 and 82 °C, respectively). On cooling from the isotropic phase, only the glass transition and crystallization are detected. Subsequent heating and cooling scans are identical. In a single-phase system melting and crystallization temperatures are always higher than glass transition temperature, while in a microphase separated system they can be reversed as the cooperative motion of each phase is independent. The presence of a glass transition in our system at a temperature higher than the melting temperature

supports the microphase separated model advanced previously for the supermolecular dendrimers self-assembled from 12Gn-AG-COOH ($n = 2-4$)¹². Melting and crystallization are associated with the alkyl tails, and the glass transition is associated with the cooperative motion of the benzyl ether groups and the polymer backbone. Conventional crystalline polymers have their glass transition below crystallization and melting². No phase transition is detected up to the decomposition temperature of high-molecular-weight 12G2-AG-PS and 12G2-AG-PMA. Both display an anisotropic texture due to the Φ_h phase formed by their cylinders (evidenced by XRD) (Table 1). This Φ_h phase is stable up to the decomposition temperature, thus demonstrating that it is generated from rigid-rod superstructures.

X-ray diffraction results confirm the hypothesis that the shape of these polymers is determined by their DP (Table 1). At low DPs they have a spherical shape, whereas at high DP, they form a cylindrical shape. Table 1 lists the lattice dimension a (in Å) for both Cub and Φ_h phases. It also lists μ' , the number of polymer repeat units that form either a sphere (in the Cub phase) or a 3.74-Å-thick columnar cross-section (in the Φ_h phase). Table 1 shows that 12G2-AG-PS (number average molecular weight $M_n = 10,800$) and 12G2-AG-PMA ($M_n = 10,700$) have Cub phases with lattice dimensions of

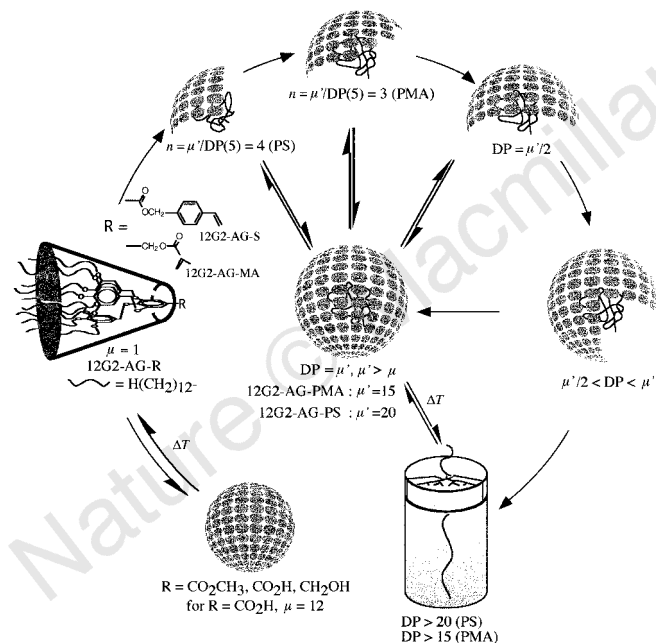


Figure 2 Mechanism of self-assembly of spherical and cylindrical supermolecules from polymer backbones jacketed with quasi-equivalent dendritic coats ($\Rightarrow$ self-assembly; $\rightarrow$ polymerization). See text for details.

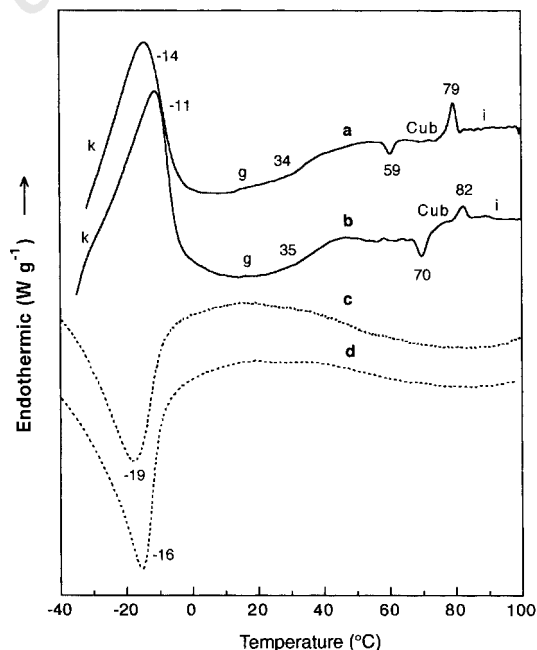


Figure 3 First heating (solid line) and cooling (dotted line) DSC traces of 12G2-AG-PS (a and c) ($M_n = 10,800$) and 12G2-AG-PMA (b and d) ($M_n = 10,700$) (k, crystal; g, glass; Cub, Cubic; i, isotropic phase).

Table 1 Characterization of spherical and cylindrical supermolecules

Compound	M_n	M_w/M_n	DP	d_{20} (g cm ⁻³)	Phase	a (Å) at T (°C)	μ' †	$n = \mu'/DP^\#$
12G2-AG-PS*	10,800‡, 12,700§	1.10	5‡	0.95	Cub	84.4(60)	20	4
12G2-AG-PS†	186,500§	2.36	84	0.96	Φ_h	47.7(85)	2	-
12G2-AG-PMA*	10,700‡, 12,300§	1.06	5‡	0.95	Cub	76.5(70)	15	3
12G2-AG-PMA†	331,900‡, 227,700§	3.11	106‡	0.96	Φ_h	48.8(85)	2	-

* Polymerization in C₆H₆ at 60 °C for 24 h, [M] = 0.16 (mol l⁻¹), [AIBN] = 0.08 (mol l⁻¹); here M is 12G2-AG-S or 12G2-AG-MA and AIBN is 2,2'-azobisisobutyronitrile.

† Polymerization in bulk at 90 °C for 4 min, [M] = 0.44 (mol l⁻¹), [AIBN] = 0.14 (mol l⁻¹).

‡ By light scattering.

§ By gel permeation chromatography (GPC) with polystyrene standards in THF; M_w is weight average molecular weight, M_n is number average molecular weight, and M_w/M_n is polydispersity.

|| Lattice parameter obtained from XRD data; for Φ_h : $a = (d_{100}^2 + d_{110}^2 + d_{200}^2)/3$ (12G2-AG-PS: $d_{100} = 41.5$, $d_{110} = 23.9$, $d_{200} = 20.6$; 12G2-AG-PMA: $d_{100} = 41.7$, $d_{110} = 24.8$, $d_{200} = 20.9$); for Cub: $a = (\sqrt{d_{200}^2 + d_{210}^2} + \sqrt{d_{211}^2 + d_{220}^2} + \sqrt{d_{310}^2 + d_{320}^2} + \sqrt{d_{321}^2 + d_{330}^2} + \sqrt{d_{331}^2 + d_{340}^2})/8$ (12G2-AG-PS: $d_{200} = 41.9$, $d_{210} = 37.7$, $d_{211} = 34.6$, $d_{220} = 29.9$, $d_{310} = 26.7$, $d_{320} = 23.4$, $d_{321} = 22.6$, $d_{330} = 21.1$; 12G2-AG-PMA: $d_{200} = 38.3$, $d_{210} = 34.2$, $d_{211} = 31.2$).

† μ' is the number of polymer repeat units containing monodendrons per spherical dendrimer or per 3.74-Å stratum of the cylindrical dendrimer.

n is the number of polymers per spherical supramolecular dendrimer.

84.4 and 76.5 Å, respectively; the Cub phase formed through the self-assembly of 12 individual 12G2-AG-COOH monodendrons¹², in contrast, has a lattice dimension of only 68.3 Å. As the monodendron units used in the three systems are of the same size, the difference in a values can only be attributed to the need to accommodate the PS and PMA backbones of 12G2-AG-PS and 12G2-AG-PMA in the centre of their spherical nanoarchitecture. As might be expected, the number of monodendrons required to form a single sphere increases from $\mu = 12$ for 12G2-AG-COOH to $\mu' = 20$ for 12G2-AG-PS and $\mu' = 15$ for 12G2-AG-PMA. (Here μ gives the number of free monodendrons that make up a sphere; μ' , in contrast, gives the number of monodendritic units that make up a sphere while being attached to a polymeric backbone.) As the two polymers have, on average, five monodendritic side-groups attached to each backbone, the formation of a single sphere requires four polymer chains (n in Table 1) in the case of 12G2-AG-PS, and three in the case of 12G2-AG-PMA.

In the case of high-DP 12G2-AG-PS and 12G2-AG-PMA, the number of monodendrons which form a 3.74-Å-thick slice of column, is $\mu' = 2$. The spacing between adjacent side-groups along an extended backbone is ~ 2 Å. Therefore, the backbone of these two polymers must adopt an almost fully extended conformation, penetrating through the centre of the cylinder (Fig. 2, bottom right).

Monolayers of the spherical and cylindrical polymers were visualized in their isotropic phase by scanning force microscopy (SFM) (Fig. 4)²². For both cylindrical and spherical 12G2-AG-PS, the monolayer thickness and the lateral diameter were 50 ± 5 Å. This value agrees with the data from XRD (Table 1). Topographic contrast (Fig. 4) shows strong molecular segregation and weak interpenetration of 12G2-AG-PS that are consistent with a high steric crowding of the dendritic coats. The geometrical constraint of the flat and hard substrate, and the cohesive forces acting during film formation from dilute solution, cause some orientational ordering of the cylindrical molecules. The average length of the straight segment (persistence length) is ~ 400 Å for 12G2-AG-PS and 600 Å for 12G2-AG-PMA (both with high DPs). Hairpin folding^{3,5}, expected for relatively long worm-like molecules, is also visible (Fig. 4b).

The mechanism of construction of spherical and cylindrical polymer backbones surrounded with quasi-equivalent dendritic coats is outlined in Fig. 2. Twelve ($\mu = 12$) 12G2-AG-COOH monodendrons self-assemble into a spherical supramolecular dendrimer which in turn self-organizes in a $Pm\bar{3}n$ cubic lattice ($a = 68.3$ Å)¹². 12G2-AG-PS with $M_n = 10,800$ (DP = 5, where DP gives the average number of side-groups attached to the polymer backbone) has a shape which is equivalent to a quarter of a sphere as four of these polymers self-assemble into a sphere ($a = 84.4$ Å) that contains a total of 20 monodendron units ($\mu' = 20$). 12G2-AG-PMA with $M_n = 10,700$ (DP = 5) has a shape equivalent to one-third of a sphere as three of these polymers are required to produce a sphere ($a = 76.5$ Å) containing 15 monodendrons ($\mu' = 15$). When $DP = \mu'/2$, the macromolecule contains half of the monodendritic units required to form a sphere; its shape is therefore likely to be hemispherical, with two polymers self-assembling into a full sphere. At first sight, difficulties are expected with this self-assembly mechanism when the polymer's DP is larger than $\mu'/2$ and slightly smaller than μ' (right side of Fig. 2). In such a case, highly dissimilar fragments (for example, a three-quarter sphere and a one-quarter sphere) generated from completely different DPs are required to produce a single sphere. However, the polydisperse molecular-weight distribution at low DPs leads to the presence of dissimilar sphere fragments in a single polymer sample thereby ensuring the ability of low-DP polymers with identical or even different backbones to self-assemble into spheres (unpublished results). When $DP = \mu'$, polydispersity effects prevent the polymer chain from assuming the expected spherical shape because the highly dissimilar

but complementary sphere fragments are not available in the same reaction product. When approaching and exceeding $DP = \mu'$, the polymer therefore has to adopt an alternative and less favourable option for the minimization of its free energy: the polymer assumes a cylindrical shape with a backbone that penetrates through its centre and the quasi-equivalent monodendrons became flat-tapered, rather than conical (bottom right in Fig. 2). At DP much larger than μ' , this is the only mechanistic option available. In the range of DPs where the polymer changes shape, the coexistence of spherical and cylindrical polymers generates a biphasic region whose width is determined by monodendron architecture, the nature of the polymer backbone and polydispersity. The change in shape from spherical to cylindrical by increasing DP demonstrates that the 12G2-AG monodendron can change its shape from the conical to flat-tapered, and therefore acts as a quasi-equivalent

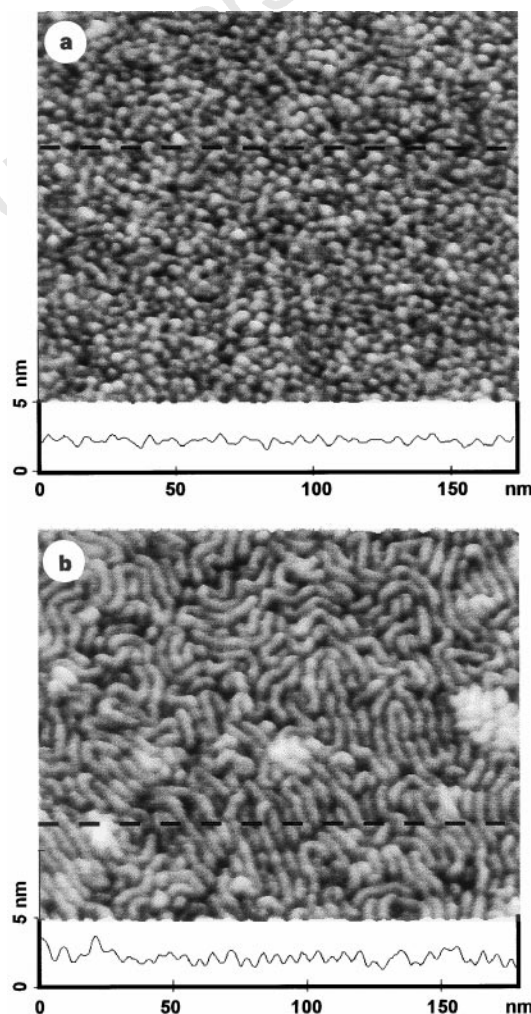


Figure 4 Topographic scanning force microscopy (SFM) images (tapping mode) of monomolecular films of 12G2-AG-PS. **a**, $M_n = 10,800$; **b**, $M_n = 186,500$. The two-dimensional profile was recorded along the dashed line within the image. Si tips with a resonance frequency of ~ 300 kHz and a probe apex with a radius of < 100 Å were used at 25 °C. Monolayers of 12G2-AG-PS on a mica substrate were obtained by spin coating at 2,000 r.p.m. from 0.2 mg ml⁻¹ chloroform solutions for high molecular weight, and from cyclohexane solutions for low molecular weight. The less-polar cyclohexane is a better solvent for the alkyl tails of the monodendrons and favoured the self-assembly of spherical structures in dilute solution. The monolayer structure of the film, coupled with the molecular resolution achieved, allows calculation of the thickness and the width of the superstructures.

building block for the formation of fundamentally different polymer shapes.

The mechanism outlined in Fig. 2 allows the rational design of polymers with well defined spherical and cylindrical shapes. The self-assembly of the dendritic coats controls the conformation of their polymer backbone such that an increase in DP changes the backbone conformation and polymer shape, respectively, from random-coiled and spherical to extended and cylindrical. The general design concept described here has been successfully tested on a library of quasi-equivalent dendritic monomers with different architecture, generation number and polymer backbone (unpublished results). Unlike conventional polymers, the physical properties of the polymers reported here are determined not only by their DP but also by their shape. Examples of extension of flexible random-coil polymer backbones were previously known only for a few natural (that is, glycoprotein)²³, and synthetic (that is, poly(macromonomer)²⁴) 'bottlebrush' polymers. Although a small change in backbone conformation occurs at the transition from nematic to smectic phase in side chain LC polymers²⁵, we are not aware of any natural or synthetic macromolecular systems that exhibit an equally dramatic and controlled change in shape and backbone conformation as a function of molecular weight as seen in this study^{10,13,24,26}. □

Received 30 May; accepted 15 October 1997.

1. Flory, P. J. *Statistical Mechanics of Chain Molecules* (Hanser, New York, 1989).
2. Keller, A. in *Sir Charles Frank, OBE, FRS, An Eightieth Birthday Tribute* 265–306 (Hilger, New York, 1991).
3. de Gennes, P. G. in *Polymer Liquid Crystals* (eds Ciferri, A., Krigbaum, W. R. & Meyer, R. B.) 115–131 (Academic, New York, 1982).
4. Ciferri, A. (ed.) *Liquid Crystallinity in Polymers* (VCH, Weinheim, 1991).
5. Hardouin, F. et al. SANS study of a semiflexible main chain liquid crystalline polyether. *Macromolecules* **28**, 5427–5433 (1995).
6. Lehn, J. M. *Supramolecular Chemistry* (VCH, Weinheim, 1995).
7. Newkome, G. R., Moorefield, C. N. & Vögtle, F. *Dendritic Molecules. Concepts, Synthesis, Perspectives* (VCH, Weinheim, 1996).
8. de Gennes, P. G. Soft matter (Nobel Lecture). *Angew. Chem. Int. Edn. Engl.* **31**, 842–845 (1992).
9. Bissell, R. & Boden, N. Liquid crystals display new potential. *Chem. Brit.* **31**, 38–41 (1995).
10. Fréchet, J. M. J. Functional polymers and dendrimers: reactivity, molecular architecture, and interfacial energy. *Science* **263**, 1710–1715 (1994).
11. Percec, V., Johansson, G., Ungar, G. & Zhou, J. Fluorophobic effect induces the self-assembly of semifluorinated tapered monodendrons containing crown-ethers into supramolecular columnar dendrimers which exhibit a homeotropic hexagonal columnar liquid crystalline phase. *J. Am. Chem. Soc.* **118**, 9855–9866 (1996).
12. Balagurusamy, V. S. K., Ungar, G., Percec, V. & Johansson, G. Rational design of the first spherical supramolecular dendrimers self-organized in a novel thermotropic cubic liquid-crystalline phase and the determination of their shape by X-ray analysis. *J. Am. Chem. Soc.* **119**, 1539–1555 (1997).
13. Tomalia, D. A. & Durst, H. D. in *Topics in Current Chemistry* Vol. 165 (ed. Weber, E.) 193–313 (Springer, Berlin, 1993).
14. Amabilino, D. B. et al. Self-assembly of [n] rotaxanes bearing dendritic stoppers. *J. Am. Chem. Soc.* **118**, 12012–12020 (1996).
15. Percec, V., Chu, P., Ungar, G. & Zhou, J. Rational design of the first nonspherical dendrimer which displays calamitic nematic and smectic thermotropic liquid crystalline phases. *J. Am. Chem. Soc.* **117**, 11441–11454 (1995).
16. Casper, D. L. D. Movement and self-control in protein assemblies. Quasi-equivalence revisited. *Biophys. J.* **32**, 103–138 (1980).
17. Levine, A. J. *Viruses* (Freeman, New York, 1992).
18. Klug, A. From macromolecules to biological assemblies. *Angew. Chem. Int. Edn. Engl.* **22**, 565–582 (1983).
19. Ackermann, A. W. & Berthiaume, L. *Atlas of Virus Diagrams* (CRC, New York, 1995).
20. Hudson, S. D. et al. Direct visualization of individual cylindrical and spherical supramolecular dendrimers. *Science* **278**, 449–452 (1997).
21. Kwon, Y. K., Chvalun, S. N., Blackwell, J., Percec, V. & Heck, J. A. Effect of temperature on the supramolecular tubular structure in oriented fibers of a poly(methacrylate) with tapered side groups. *Macromolecules* **28**, 1552–1558 (1995).
22. Sheiko, S. S., Gauthier, M. & Möller, M. Monomolecular films of arborescent polystyrene. *Macromolecules* **30**, 2343–2349 (1997).
23. Wainwright, S. A. in *Biomimetics. Design and Processing of Materials* (eds Sarikaya, M. & Aksay, I. A.) 1–12 (AIP Press, New York, 1995).
24. Wintermantel, M. et al. Lyotropic phases formed by molecular bottlebrushes. *Angew. Chem. Int. Edn. Engl.* **34**, 1472–1474 (1995).
25. Hardouin, F. Mesomorphic properties and SANS conformational study in different liquid crystalline polymers. *Macromol. Symp.* **98**, 329–339 (1995).
26. Thomas, E. L., Anderson, D. M., Henkee, C. S. & Hoffman, D. Periodic area-minimizing surfaces in block copolymers. *Nature* **334**, 598–601 (1988).
27. Percec, V., Ahn, C.-H. & Barboiu, B. Self-encapsulation, acceleration and control in the radical polymerization of monodendritic monomers via self-assembly. *J. Am. Chem. Soc.* **119**, 12978–12979 (1997).

Acknowledgements. This work was supported by the US NSF, the UK Engineering and Physical Science Research Council, the Synchrotron Radiation Source at Daresbury (beamtime and technical assistance), the Deutsche Forschungsgemeinschaft and a Humboldt research award (to V.P.).

Correspondence and requests for materials should be addressed to V.P. (e-mail: vxp5@po.cwru.edu).

Sulphur-radical control on petroleum formation rates

Michael D. Lewan

US Geological Survey, Box 25046, MS 977, Denver Federal Center, Denver, Colorado 80225, USA

Most petroleum is formed through the partial decomposition of kerogen (an insoluble sedimentary organic material) in response to thermal stress during subsurface burial in a sedimentary basin^{1,2}. Knowing the mechanisms and kinetics of this process allows the determination of the extent and timing of petroleum formation, which, in turn, are critical for evaluating the potential for petroleum occurrences within a sedimentary basin. Kinetic models of petroleum generation are derived mainly from pyrolysis experiments^{1,2}, in which it is usually assumed that formation rates are controlled by the strength of the bonds within the precursor compounds: this agrees with the observation that petroleum formation rates increase with increasing sulphur content of thermally immature kerogen^{2–4}, C–S bonds being weaker than C–C bonds. However, this explanation fails to account for the overall composition of petroleum. Here I argue, on the basis of pyrolysis experiments, that it is the presence of sulphur radicals, rather than the relative weakness of C–S bonds, that controls petroleum formation rates. My findings suggest that the rate of petroleum formation depends critically on the concentration of sulphur radicals generated during the initial stages of thermal maturation. The proposed mechanism appears to provide a realistic explanation for both the overall composition of petroleum and the observed variation in formation rates.

Kinetic controls on petroleum generation require pyrolysis experiments to be conducted at higher temperatures (300 to 650 °C) than those experienced in sedimentary basins (70 to 200 °C) if experiments are to be completed within reasonable durations of hours to weeks, as opposed to geological durations of millions of years. This difference in timing and temperature means that results from laboratory pyrolysis experiments must be extrapolated to lower temperatures and longer times. Extrapolations typically employ the Arrhenius expression with experimentally derived activation energies (E_a) and frequency factors (A_0). The plot in Fig. 1 shows the relationship between organic-sulphur mole fraction of type-II kerogens (the most common source of petroleum) and their activation energy for petroleum formation as determined by the amount of expelled oil generated during hydrous pyrolysis⁵. These activation energies and their corresponding frequency factors give reasonable extrapolations for petroleum formation in sedimentary basins and show geologically significant variations in the timing and thermal-stress levels for petroleum formation⁶. Experimental relationships like the one shown in Fig. 1 and natural subsurface data suggesting petroleum formation at abnormally low thermal-stress levels have led to the interpretation that rates of petroleum generation are faster in high-sulphur type-II kerogens because C–S bonds cleave more readily than C–C bonds^{2–4,7–10}. Although covalent C–S bonds have bond dissociation energies that can be 36 kJ mol^{–1} less than comparable C–C bonds¹¹, activation energies in Fig. 1 represent a continuum that extends over a range of 100 kJ mol^{–1}. In addition, molecular modelling indicates that covalent sulphide linkages do not have a significant effect on the bond strength of adjacent carbon–carbon bonds, which have bond strengths that do not vary by more than 15 kJ mol^{–1} depending on their proximity to carbon atoms with primary and tertiary hydrogens^{12,13}. Therefore, it is difficult to explain how alkyl C–C bonds can be cleaved to generate the multitude of organic components in petroleum at lower thermal-stress levels as a result of the cleavage of weaker and limited C–S bonds in type-II kerogen.

Table 1 Initial reactants and recovered products

	Initial DEDS (mmoles)			
	0.000	0.982	1.767	3.370
Initial PDD	36.796	36.646	36.548	36.548
Recovered liquid	35.197	35.298	34.905	35.181
Recovered gas	0.416	1.094	1.484	2.993
H ₂ S	0.000	0.206	0.487	1.189
C ₂ H ₆	0.087	0.319	0.648	1.448
C ₂ H ₄	0.000	0.000	0.001	0.000
CH ₄	0.071	0.249	0.197	0.210
C ₃ H ₈	0.023	0.022	0.041	0.063
C ₃ H ₆	0.000	0.000	0.009	0.010
H ₂ S/C ₂ H ₆ *	—	0.888	0.868	0.874

Results of experiments with 1-phenyldodecane (PDD) and different concentrations of diethyldisulphide (DEDS) at 350 °C for 72 h. All quantities are in mmol.

* 0.087 mmol of ethane generated in the experiment with no DEDS was subtracted from the ethane quantities used to calculate this ratio.

The alternative mechanism suggested here emphasizes the concentration of sulphur radicals, rather than bond strengths, as the rate-controlling factor in petroleum formation. In this mechanism, the greater the concentration of initiating sulphur radicals formed during thermal maturation of sedimentary organic matter, the faster the rate of thermal cracking of the organic matter to petroleum. The wide range of carbon numbers in *n*-alkane distributions and limited abundance of iso-alkanes in petroleum indicate that most of the detectable components in petroleum are a result of a free-radical mechanism^{14–16}. Naturally matured and experimentally pyrolysed kerogens have been shown by electron spin resonance spectroscopy to develop free radicals that increase in concentration with increasing thermal stress¹⁷. In the aliphatic domains of a kerogen, the abstraction of a hydrogen atom from an alkyl group initiates thermal cracking of C–C bonds, which subsequently leads to a chain of thermal-cracking reactions. This abstraction may occur through the interaction of a pre-existing radical derived from the cleavage of weaker bonds within the

kerogen. C–S and S–S bonds are likely sources of these initiating free radicals, with the latter being particularly favourable for generating thiyl radicals at low thermal-stress levels^{18,19}. As a result, rates of thermal cracking are expected to increase in proportion to the number of C–S and S–S bonds in the original thermally immature type-II kerogen. By this mechanism, cleavage of the weaker C–S and S–S bonds is not responsible for generation of each of the many constituents of petroleum, but rather provides the initiating free radicals that allow the thermal cracking of C–C bonds to occur at a faster rate and lower thermal stress.

I have tested this proposed mechanism with a series of model-compound experiments conducted with 1-phenyldodecane (PDD) and varying concentrations of diethyldisulphide (DEDS). Long-chain alkyl-substituted aromatic compounds such as PDD have been proposed as relevant model compounds for studying the thermal behaviour of complex natural organic materials including coals, heavy oils and asphaltene²⁰. Thermal decomposition of PDD has been extensively studied and shown to occur through a free radical mechanism²¹ that can be described by a first-order rate function at temperatures ranging from 350 to 400 °C (ref. 22). Although organic sulphur occurs in a variety of moieties in type-II kerogen (for example, thiophene, sulphoxide, sulphides)²³, homolytic cleavage of sulphides, especially disulphides, is a likely source of sulphur radicals (thiyls) at low thermal-stress levels²⁴. DEDS was used in these experiments as a source of initiating sulphur radicals because the potential products that it can form directly (H₂S, ethane and ethanethiol) are readily distinguishable for the thermal decomposition products of PDD. All of the experiments were conducted in 22-cm³ stainless steel 316 reactors, with nine grams of 97% PDD and varying amounts (0 to 412 mg) of 99% DEDS (Table 1). The reactors were heated at 350 °C for 72 hours in a fluidized sand bath.

As in previous pyrolysis studies on PDD^{21,22}, the main cracking products in the recovered liquid were toluene, undecenes, ethylbenzene, and decane. Only traces of recombination products were

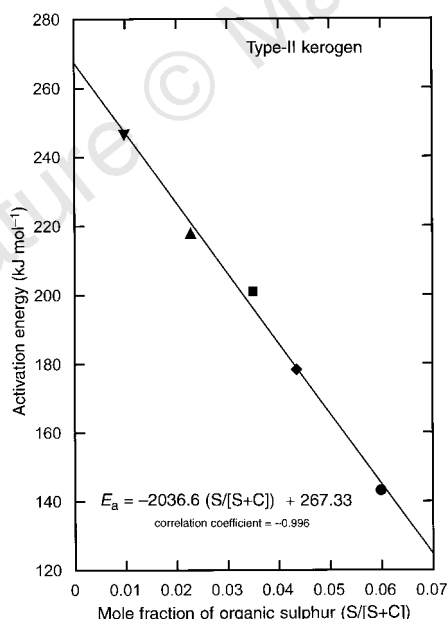


Figure 1 Plot of the activation energy for expelled oil generation from type-II kerogen against the sulphur mole fraction ($S/(S+C)$) of the original thermally immature kerogen; New Albany Shale (▼), Woodford Shale (▲), Alum Shale (■), Phosphoria Formation (◆), and Monterey Formation (●). Sulphur values are for organic sulphur and do not include inorganic sulphur (such as pyrite/marcasite). All the activation energies and their frequency factors are reported in ref. 6, except those for the New Albany Shale, which were determined by J. Guthrie and M.D.L. to be 247.7 kJ mol⁻¹ and 2.315×10^{15} s⁻¹, respectively.

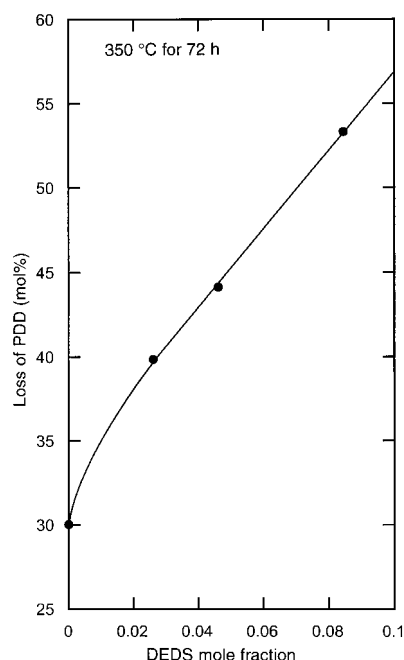


Figure 2 Plot of mol% of 1-phenyldodecane (PDD) degraded against diethyldisulphide (DEDS) mole fraction ($DEDS/(DEDS + PDD)$) used in the experiments at 350 °C for 72 h. Loss of PDD was determined by gas chromatography with a flame ionization detector calibrated with external standards. Note that experiments were carried out under anhydrous conditions in order to focus on the significance of sulphur radicals on the free-radical mechanism. (H₂O plays a part in quenching free radical sites and can reduce the rate of thermal cracking reactions¹⁶.)

detected with 1-undecyl-dodecyl benzene being the most apparent. None of the original DEDS was detected in the liquid product, but ethane and hydrogen sulphide were the main components in the recovered gases (Table 1). No organic sulphur compounds were detected in the liquid product by gas chromatography with a flame photometric detector. The unaccounted for sulphur left can be attributed to H_2S , which has a high solubility in liquid hydrocarbons^{25,26} and can react with stainless steel 316 to form a pyrrhotite ($Fe_{1-x}S$) film on reactor walls⁵. Although the liquid products were proportionally the same in all the experiments, the amount of reaction as determined by the amount of recovered PDD was considerably different. As shown in Fig. 2, the extent of the overall reaction, as measured by the loss of PDD, increased from 30 mol% in the experiment with no DEDS to 53 mol% in the experiment with 3.37 mmol DEDS. This significant increase in the cleavage of the same carbon-carbon bonds at the same thermal stress level demonstrates the importance of the initiating sulphur radicals in thermal cracking.

In addition to providing a more sound mechanism to explain variations in the kinetics of petroleum formation, this mechanism also has implications for natural gas formation through thermal cracking of migrated petroleum in reservoirs. Once petroleum has been expelled from a source rock it will no longer reside in the initiating free-radical field in which it was generated. As a result, generation of natural gas by further cracking of C-C bonds in expelled petroleum will occur at slower rates. The intercept of the relationship in Fig. 1 suggests that the highest activation energy for petroleum formation from type-II kerogen containing no initiating sulphur radicals is 271 kJ mol^{-1} , which falls well within the $240\text{--}293 \text{ kJ mol}^{-1}$ range of activation energies ascribed to petroleum cracking^{27,28}.

Another important implication is that laboratory pyrolysis methods used to derive kinetic parameters for petroleum formation in sedimentary basins must simulated this mechanism in order to provide meaningful extrapolations to the lower temperatures and longer times experienced in the subsurface of sedimentary basins. In particular, the mechanism is not likely to be stimulated by open-system pyrolysis methods² used to determine kinetic parameters, in which volatilized products are collected by sweeping carrier gas over a sample as it is isothermally heated from 300 to 600 °C within tens of minutes. The problem here is that initiating sulphur radicals generated in the early stages of heating in open-system pyrolysis are removed from the maturing organic matter and unavailable to participate fully in initiating thermal cracking reactions. Conversely, closed-system pyrolysis, such as hydrous pyrolysis⁵, at temperatures between 280 and 360 °C would maintain initiating sulphur radicals in contact with thermally maturing kerogen. This difference between open- and closed-system pyrolysis provides an explanation as to why the former does not always show a relationship between rates of petroleum generation and organic sulphur content of type-II kerogen²⁹. □

Received 9 May; accepted 30 September 1997.

- Hunt, J. M. *Petroleum Geochemistry and Geology* 2nd Edn (Freeman, New York, 1996).
- Tissot, B. P., Pelet, R. & Ungerer, P. Thermal history of sedimentary basins, maturation indices, and kinetics of oil and gas generation. *Am. Assoc. Petrol. Geol. Bull.* **71**, 1445–1466 (1987).
- Lewan, M. D. Evaluation of petroleum generation by hydrous pyrolysis. *Phil. Trans. R. Soc. Lond. A* **315**, 123–134 (1985).
- Tomic, J., Behar, F., Vandenbroucke, M. & Tang, Y. Artificial maturation of Monterey kerogen (Type II-S) in a closed system and comparison with Type II kerogen: implications on the fate of sulfur. *Org. Geochem.* **23**, 647–660 (1995).
- Lewan, M. D. Laboratory simulation of petroleum formation: hydrous pyrolysis, in *Organic Geochemistry* (eds Engel, M. H. & Macko, S. A.) 419–442 (Plenum, New York, 1993).
- Hunt, J. M., Lewan, M. D. & Hennen, R. J.-C. Modeling oil generation with time-temperature index graphs based on the Arrhenius equation. *Am. Assoc. Petrol. Geol. Bull.* **75**, 795–807 (1991).
- Baskin, D. K. & Peters, K. E. Early generation characteristics of a sulfur-rich Monterey kerogen. *Am. Assoc. Petrol. Geol. Bull.* **76**, 1–13 (1992).
- Orr, W. L. Kerogen/asphaltene/sulfur relationships in sulfur-rich Monterey oils. *Org. Geochem.* **10**, 499–516 (1986).
- Tannenbaum, E. & Aizenshtat, Z. Formation of immature asphalt from organic-rich carbonate rocks—I. Geochemical correlation. *Org. Geochem.* **8**, 181–192 (1985).
- Petersen, N. F. & Hickey, P. J. California Plio-Miocene oils: evidence of early generation, in *Exploration for Heavy Oil and Natural Bitumen* (ed. Meyer, R. E.) 351–359 (Am. Assoc. Petrol. Geol., Studies in Geology no. 25, 1987).

- McMillen, D. F. & Golden, D. M. Hydrocarbon bond dissociation energies. *Ann. Rev. Phys. Chem.* **33**, 493–532 (1982).
- Claxton, M. J., Patience, R. L. & Park, P. J. D. Molecular modelling of bond energies in potential kerogen sub-units, in *Poster Sessions from the 16th International Meeting on Organic Geochemistry*, Stavanger, Norway (ed. Oygard, K.) 189–201 (Falch Hurtigtrykk, Oslo, 1993).
- Patience, R. L. & Claxton, M. J. Bond energies in alkyl-substituted organic sulfur compounds, in *Poster Sessions from the 16th International Meeting on Organic Geochemistry*, Stavanger, Norway (ed. Oygard, K.) 217–219 (Falch Hurtigtrykk, Oslo, 1993).
- Eisma, E. & Jurg, J. W. Fundamental aspects of the generation of petroleum, in *Organic Geochemistry: Methods and Results* (eds Eglinton, G. & Murphy, M. T. J.) 676–698 (Springer, New York, 1969).
- Beltrame, P. L., Carniti, P., Audisio, G. & Bertini, F. Catalytic degradation of polymers: part II—degradation of polyethylene. *Polymer Degradation and Stability* **26**, 209–220 (1989).
- Lewan, M. D. Experiments on the role of water in petroleum formation. *Geochim. Cosmochim. Acta* **61**, 3691–3723 (1997).
- Bakr, M., Akiyama, M. & Sanadia, Y. In situ high temperature ESR measurements for kerogen maturation. *Org. Geochem.* **17**, 321–328 (1991).
- Field, L., in *Organic Chemistry of Sulfur*, ch. 7 (ed. Oae, S.) 303–382 (Plenum, New York, 1977).
- Clark, P. D. & Oriakhi, C. O. Studies on the stability of organic disulfide–polysulfide sulfur solvents used in sour gas production. *Energy Fuels* **6**, 474–477 (1992).
- Savage, P. E., Jacobs, G. E. & Javanmardian, M. Autocatalysis and aryl–alkyl bond cleavage in 1-dodecylpyrene pyrolysis. *Ind. Eng. Chem. Res.* **28**, 645–654 (1989).
- Savage, P. E. & Klein, M. T. Discrimination between molecular and free-radical models of 1-phenyldodecane pyrolysis. *Int. Eng. Chem. Res.* **26**, 374–376 (1987).
- Blouri, B., Hamdan, F. & Herauld, D. Mild cracking of high-molecular-weight hydrocarbons. *Ind. Eng. Chem. Res.* **24**, 30–37 (1985).
- Nelson, B. C. et al. Transformations in organic sulfur speciation during maturation of Monterey Shale: constraints from laboratory experiments, in *Geochemical Transformations of Sedimentary Sulfur* (eds Vairavamuthy, M. A. & Schoonen, M. A. A.) 138–166 (Am. Chem. Soc. Symp. Series 612, Washington DC, 1995).
- Pryor, W. A. *Mechanisms of Sulfur Reactions* (McGraw-Hill, New York, 1962).
- Tremper, K. K. & Prausnitz, J. M. Solubility of inorganic gases in high-boiling hydrocarbon solvents. *J. Chem. Eng. Data* **21**, 295–299 (1976).
- Yokoyama, C., Usui, A. & Takahashi, S. Solubility of hydrogen sulfide in isooctane, n-decane, n-tridecane, n-hexadecane, and squalane at temperatures from 323 to 523 K and pressures up to 1.6 MPa. *Fluid Phase Equilibria* **85**, 257–269 (1993).
- Horsfield, B. H., Schenk, H. J., Mills, N. & Welte, D. H. An investigation of the in-reservoir conversion of oil to gas: compositional and kinetic findings for closed-system programmed-temperature pyrolysis. *Org. Geochem.* **19**, 191–204 (1991).
- Pepper, A. S. & Dodd, T. A. Simple kinetic models of petroleum formation. Part II: oil-gas cracking. *Mar. Petrol. Geol.* **12**, 321–340 (1995).
- Reynolds, J. G., Burnham, A. K. & Mitchell, T. O. Kinetic analysis of California petroleum source rocks by programmed temperature microanalysis. *Org. Geochem.* **23**, 109–120 (1995).

Acknowledgements. I gratefully acknowledge reviews by H. Helgeson and S. Larter, and helpful suggestions by Z. Aizenshtat, J. Clayton, R. Dias, D. King, P. Lillis, T. Ruble and K. Varnes.

Correspondence should be addressed to M.D.L. (e-mail: mlewan@usgs.gov).

Constraints on Earth accretion deduced from noble metals in the oceanic mantle

Jonathan E. Snow* & Gerhard Schmidt†

* Max-Planck-Institut für Chemie (Otto-Hahn-Institut) Postfach 3060, 55020 Mainz, Germany

† Institut für Kernchemie, Johannes Gutenberg-Universität Mainz, Fritz-Straßmann-Weg 2, 55099 Mainz, Germany

If the Earth's mantle were in equilibrium with its core, the mantle would contain three orders of magnitude less of the noble metals (platinum-group elements Pt, Os, Ir, Ru, Pd and Rh, plus Au and Re) than are observed. An explanation put forward to account for this disparity has been that the last 1% of the Earth's accretion occurred after the iron-rich core had separated from the mantle^{1,2}. Recent debate has accordingly centred on which meteorite class or classes made up this 'late veneer' of accretion³. Here we present analyses of noble-metal concentrations in oceanic peridotites (plutonic rocks which are thought to represent samples of the Earth's upper mantle). We find that the average oceanic-mantle Os/Ir ratio is indistinguishable from that in the CI-type carbonaceous chondrites⁴, but that Ru/Ir, Pt/Ir, Rh/Ir and Pd/Ir ratios are about 40% higher. A late veneer composed of strictly CI-type carbonaceous chondritic composition is therefore not compatible with these observations. The data also allows us to rule out other carbonaceous chondrites⁵, enstatite chondrites^{6,7} and ordinary chondrites⁸ as significant late veneer components. We propose that mixing of differentiated outer-core material back into the

mantle after core separation could account for the observed noble-metal ratios and abundances in the mantle without any late accretionary veneer.

Abyssal peridotites are fragments of the oceanic upper mantle that crop out on the sea floor through tectonic processes. They provide a particularly good means of testing the predicted noble-metal abundances of different mantle evolution paths because they are younger than typical subcontinental lithospheric mantle and have been through a less complex lithospheric and crustal history. For this reason we have measured noble-metal abundances in a suite of nine abyssal peridotites from the Atlantic⁹, Pacific¹⁰ and Indian oceans^{11,12} (Table 1). In Fig. 1a the data are plotted normalized to Ir and carbonaceous chondrites (C-chondrites)⁴. The remaining elements are also broadly chondritic, but there is a 30–50% average enrichment of the elements Ru, Pt, Rh and Pd with respect to Ir and Os.

All of the samples are melting residues, and this process can change noble-metal concentrations. Re and Au behave in an incompatible manner and are extracted from such residues (see Fig. 1a). The other noble metals are highly compatible in the

presence of residual sulphide^{13,14}. Almost all abyssal peridotites show some evidence for primary sulphide, and there is no evidence from mid-ocean ridges that melting there proceeds to the point where all sulphide has been consumed. Noble-metal ratios are thus buffered during melting by the residual sulphide phase, although their absolute concentrations may vary by as much as 20% owing to the extraction of silicate melt.

The noble metals are stable during oceanic serpentinization because this is among the most reducing environments on the Earth (free troilite and native iron are common occurrences¹⁶), and the noble metals are all extremely unreactive under such conditions. Oceanic serpentinization does change the absolute concentrations of noble metals because of the formation of low-density serpentine, but noble-metal ratios are once again not affected. Weathering, as distinct from serpentinization, can also affect peridotite compositions. The pair of samples AII107:61–83C and 61–83W are a fresh and a weathered sample, respectively, from core and rim of a single fine-grained sample. The weathered sample shows strong weathering effects on major and trace elements^{11,12}. All the noble metals except for Re are slightly enriched during this weathering (as are Cr

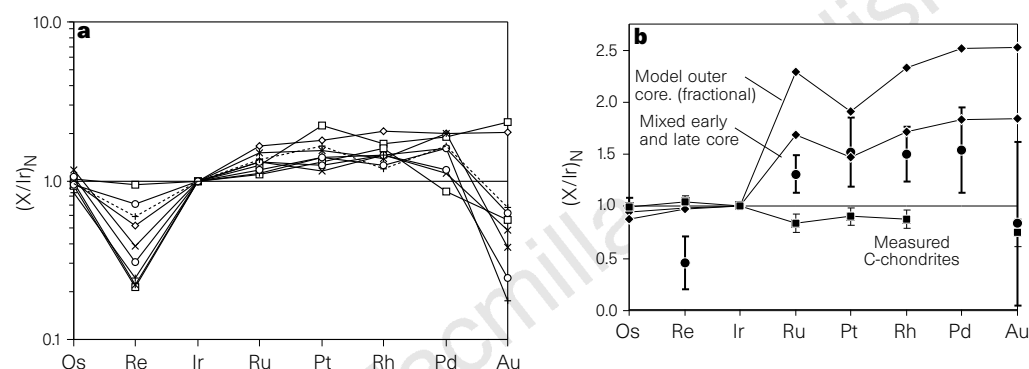


Figure 1 Abyssal peridotite noble metal abundances. **a**, Data normalized to Ir and C-chondrites⁴. **b**, Abyssal peridotite averages (closed circles). Error bars are population standard deviations (σ). Shown for comparison are recently measured carbonaceous chondrites (closed squares)⁵, a core crystallization model, and the same core model mixed with 40% of chondritic early Archaean outer core to form a mantle with present-day noble-metal ratios (diamonds).

Table 1 Noble metals determined in oceanic peridotites

Sample	Type	Os	Re	Ir	Ru	Pt	Rh	Pd	Au
AII107:61–83 W	Lh	1.88 ± 04	0.08 ± 1	1.83 ± 1	4.73 ± 16	7.11 ± 20	1.15 ± 10	4.42 ± 20	1.23 ± 2
AII107:61–83 C	Lh	1.52 ± 06	0.11 ± 1	1.39 ± 1	2.68 ± 12	6.70 ± 20	0.72 ± 12	3.16 ± 10	1.07 ± 1
AII107:40–35	Hx	2.21 ± 14	0.04 ± 1	2.25 ± 1	3.86 ± 27	6.40 ± 40	1.09 ± 11	2.33 ± 10	0.42 ± 1
PS86:6–37	Lh	3.06 ± 06	0.08 ± 1	2.47 ± 1	5.04 ± 22	6.14 ± 30	1.09 ± 08	3.31 ± 10	0.40 ± 1
ODP147 895D 4R-3:050–056 6B	Hx	3.36 ± 04	0.06 ± 1	3.26 ± 1	7.63 ± 17	10.85 ± 20	1.37 ± 10	7.79 ± 20	0.41 ± 1
ODP147 895D 2R-1:042–045 6A	Hx	2.62 ± 02	0.06 ± 1	2.35 ± 1	4.82 ± 09	6.37 ± 10	1.04 ± 09	3.34 ± 20	0.19 ± 1
ODP153 920D 13R-1:033–038 2B	Hx	3.05 ± 06	0.07 ± 1	3.44 ± 1	5.96 ± 25	10.40 ± 10	1.52 ± 11	6.66 ± 20	0.20 ± 1
ODP153 920D 5R-2 0–5 1	Hx	3.29 ± 12	0.16 ± 1	3.25 ± 1	6.85 ± 47	11.64 ± 30	1.18 ± 10	6.55 ± 20	0.73 ± 1
ODP153 920D 5R-2 45–519A	Hx	3.33 ± 05	0.20 ± 2	3.35 ± 1	6.13 ± 25	10.15 ± 50	1.27 ± 18	6.54 ± 40	0.69 ± 1
C-chondrites ⁴		486	38.3	459	714	994	140 (ref. 5)	556	152
Concentrations normalized to C-chondrites and Ir									
Average		0.98	0.46	1.00	1.31	1.52	1.50	1.54	0.83
S.D. (σ)		0.09	0.25	0.00	0.18	0.33	0.26	0.41	0.78
S.D. of the mean (σ_M)		0.03	0.09	0.00	0.06	0.11	0.09	0.14	0.28
Standard concentrations									
UMT-1 (CANMET standard)		6.0 ± 3	–	7.31 ± 10	8.2 ± 10	–	7.57 ± 11	98 ± 2	27 ± 0.5
UMT-1 (CANMET standard)		7.23 ± 3	2.99 ± 3	7.82 ± 10	9.6 ± 5	–	9.87 ± 20	93 ± 2	32.6 ± 0.2
UMT-1 (certified values)		8.0*	3.0*	8.8 ± 6	10.9 ± 15	129 ± 5	9.5 ± 11	106 ± 3	48 ± 2
UBN (CRPG serpentinite std)	Lh	1.83 ± 04	0.04 ± 1	1.83 ± 1	3.21 ± 13	8.36 ± 20	0.86 ± 08	4.40 ± 20	0.47 ± 1
UBN (CRPG serpentinite std)	Lh	1.85 ± 03	0.05 ± 1	1.89 ± 1	3.14 ± 12	15.67 ± 10	0.89 ± 10	4.60 ± 20	0.47 ± 1
Model parameters†									
Reference		25	25	25	25	26	27	26	25
Slope		–	0.904	0.877	0.064	0.243	0.5	–	–0.076
K_D		19	17.3	16.8	2.15	5.37	1.8	0.5	0.38
5.5% crystallized core (p.p.b.)		176	15	188	669	776	134	572	157
(X/Ir) _N		0.882	0.97	1	2.29	1.91	2.33	2.51	2.53

Concentrations (in p.p.b.) determined by instrumental neutron activation analysis (INAA) with NiS fire assay preconcentration²⁹. Counting errors are reported relative to the last significant digit. Unaccounted for uncertainties might include absolute variations due to the yield of the NiS fire assay, and sample inhomogeneity. We estimate the true uncertainty at 7–10% for absolute concentrations and 4–5% for inter-element ratios.

* Provisional standard values.

† Model details: Closed system fractional distribution coefficients (K_D) were determined from log-log concentration plots³³. The slopes of correlations in log-log space are related by $K_{De} = K_{Da} / (\text{Slope}_{ab} - 1) + 1$ where Slope_{ab} is the slope of $\log(a)$ to $\log(b)$. To convert these relationship into partition coefficients, one absolute partition coefficient must be known or estimated. We use a K_D for Os of 19 (ref. 24). The fractional crystallization equation is then $C_i = C_j F^{(K_D - 1)}$ (where C is concentration) using a C-chondritic starting composition⁴. In the case of Pd (ref. 26), Os and Ir data were not available, so the K_D value for each of these elements was estimated as an average of the values obtained relative to Pt and Ru (ref. 25). Walker *et al.* used only the most primitive melts (the IIA class) to calculate values of K_D , as there is a slight change in slope between early and late-stage melts, perhaps due to changing liquidus phase relations. The slopes shown above include all of class IIA, as the model calculation is not sensitive to changes in slope of that magnitude.

and Al in this rock¹¹), but noble-metal ratios are not affected. Re seems to be depleted during weathering. All other samples were carefully chosen to exclude known weathering effects^{11,12}. The drilled samples were never exposed on the ocean floor, and so are not subject to weathering. Because noble-metal concentrations are affected by melting, serpentinization and weathering but their ratios are not, all data are presented here as ratios double normalized to Ir and C-chondrites⁴.

Noble-metal data from continental mantle rocks, notably those from the Eastern Pyrenees ultramafic complex¹⁷ and Zabargad island¹⁸, show similar trends, but the Zabargad data show a much larger positive Pd anomaly. Neither this anomaly nor the variability in Ru/Ir and Rh/Ir seen in all continental data sets are present to nearly the same degree in the abyssal peridotite data. This suggests that some noble-metal anomalies in mantle rocks are due to a process that occurs locally during the formation of subcontinental lithospheric mantle but not during the formation of oceanic lithospheric mantle. Such a result is reasonable, because abyssal peridotites have undergone a much shorter lithospheric history than most orogenic theralites¹⁵ and mantle xenoliths.

This data set tightens the constraints on hypotheses that seek to explain the pattern of noble-metal variations in the mantle. If noble metals in the mantle are derived from the accretion of a late veneer, then the upper mantle should have noble-metal abundances and ratios that reflect the composition of the late-accreted material. But the abyssal peridotite data show systematic, significant differences from C-chondritic values (this is the most commonly cited candidate material) for the elemental ratios of Ru/Ir, Pt/Ir, Rh/Ir and Pd/Ir (Fig. 1b). Note that the error bars in Fig. 1b show population standard deviations (σ), the mean of the abyssal peridotite data for each of these ratios lies 4–5 standard errors of the mean (σ_M) higher than C-chondritic^{4,5}.

On the basis of noble metals, the postulated late veneer material cannot be an unmodified mixture of carbonaceous chondrites (classes CI, CV, CM and CO). This conclusion is not sensitive to the values we have chosen for normalization⁴. In comparison to other estimates of the noble-metal composition of the Solar System, the normalization values we have chosen⁴ are the *least* supportive of our conclusion. If other normalization values are used, either measured or compiled, our conclusion is strengthened^{5,19–21}.

Enstatite (E) chondrites might seem to fulfil the requirements for a late veneer material for noble metals and Os isotopes because they have slightly higher mean Pt/Ir, Re/Os and ¹⁸⁷Os/¹⁸⁸Os ratios than do C-chondrites^{3,17}. However, the mean of the C-chondrite normalized Os/Ir ratios for abyssal peridotites in Table 1 ($0.98 \pm 0.03 \sigma_M$)

is significantly lower than the mean for E-chondrites^{6,7} ($1.09 \pm 0.01 \sigma_M$, Fig. 2), although individual analyses from the two data sets do overlap. The mean Ru/Ir (shown in Fig. 2) and Pd/Ir of the enstatite chondrites are also close to C-chondritic, so these elements do not fit the abyssal peridotite data either^{6,7}. Ordinary chondrites (H, L and LL classes) have C-chondritic Os/Ir ratios (Fig. 2), but their Ru/Ir (Fig. 2), Pt/Ir and Rh/Ir ratios are still significantly lower than those we have measured⁸. Although we feel that this rules out ordinary chondrites completely as a late veneer material, if a late veneer hypothesis is to be entertained at all, ordinary chondrites must be considered the leading candidate veneer material.

The oceanic mantle noble-metal pattern shown in Fig. 1 is much more suggestive of evolved magmatic iron meteorites (classes IIAB, IIIA and IVAB, which represents asteroidal cores) than it is of chondrites. Late accretion of iron meteorites, however, cannot account for the observed fractionation—as both evolved and residual iron meteorites would be accreted, the net input from iron meteorites would still be C-chondritic. We feel that a much better interpretation of the observed patterns is that the noble-metal abundances of the mantle were modified by addition of a component with a fractionation similar to that of evolved magmatic iron meteorites. This component could have come from the Earth's outer core.

Mixing of outer-core material with mantle material has been proposed as a mechanism for the formation of high ¹⁸⁷Os/¹⁸⁸Os and ¹⁸⁶Os/¹⁸⁸Os ratios in some mantle plumes^{22–24}. The outer core is thought to be enriched in Re, Ru, Pt, Au and probably Pd with respect to Ir and Os, analogous to magmatic iron meteorites. Mixture of such material back into a C-chondritic upper mantle would reproduce the abundance patterns we have measured in the oceanic mantle. It is possible to test such a core-entrainment hypothesis quantitatively using model calculations^{22,24} for all the noble metals as well as Re and Au. Assuming a distribution coefficient $K_D (=C_{\text{solid metal}}/C_{\text{liquid metal}}$, where C is concentration) for Os of 19 (ref. 24), values of K_D can be calculated for Re, Ir, Ru, Pt, Pd and Au for the IIAB group^{25–27} (see Table 1). After 5.5% fractional crystallization of a chondritic starting composition⁴, the liquid has a composition as shown in Fig. 1b. Ru/Ir, Pt/Ir, Rh/Ir and Pd/Ir are strongly elevated with respect to C-chondrites, but Os/Ir remain essentially constant, just as observed in the abyssal peridotite data.

Outer-core entrainment by deep-seated mantle plumes thus explains the noble-metal concentrations in the mantle with no need for a late accretionary veneer. In the early Archaean era, before the onset of significant inner-core crystallization, any entrained core material would have a C-chondritic noble-metal signature. The present-day outer core would have the fractionated pattern as shown in Fig. 1b, but the present mantle would be a mixture of old, relatively unfractionated outer-core material and more recent fractionated material. This hypothesis would predict significantly more fractionated noble-metal patterns associated with mantle plumes (as for example in asthenospheric xenoliths from hotspot volcanoes) than from the depleted mantle peridotites examined here. These results provide some confirmation for the hypothesis that ¹⁸⁶Os/¹⁸⁸Os anomalies observed in hotspot-derived magmatic sulphides are also due to entrainment of core material in deep-seated mantle plumes^{22–24}.

Deep recycling of continental or oceanic material could lead, in principle, to the formation of noble-metal patterns such as the ones observed here. The noble metals are strongly fractionated from one another in both continental and oceanic crust, and in a fashion similar to our observations^{28–30}. However, the absolute concentrations of the noble metals in both reservoirs are so low that unreasonably large amounts of recycled material would be required (70% for continental crust, 50% for oceanic crust). Thus, deep recycling cannot account for the noble-metal fractionations we observe.

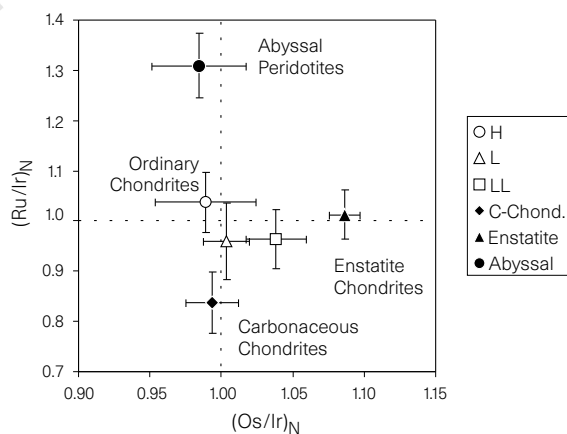


Figure 2 Distribution of Os/Ir and Ru/Ir ratios for mean abyssal peridotites (closed circles) and chondrites. Data are normalized to carbonaceous chondrites⁴. Error bars show one standard error of the mean (σ_M). Chondrite data include C-chondrites (closed diamonds)⁶, metal phases from ordinary chondrites (open symbols)⁸ and enstatite chondrites (closed triangles)^{6,34}.

A striking feature of our model is the requirement that the depleted mantle be pervasively mixed with material derived from deep-seated mantle plumes. We explain this by the fact that the noble metals are highly compatible during silicate melting. They thus remain trapped in the uppermost mantle when the plume head reaches the surface. Moreover, there is the proposal that such deep-seated mantle plumes account for all or most of the upward convective mass transfer in the mantle, and that plume head melting is primarily responsible for the formation of the depleted mantle (refs 31, 32, and J. P. Morgan and W. J. Morgan, manuscript in preparation). Although such models are currently controversial, our evidence would tend to support them. □

Received 26 February; accepted 7 October 1997.

- Chou, C. L. Fractionation of siderophile elements in the Earth's upper mantle. *Proc. Lunar Planet. Sci. Conf.* **9**, 219–230 (1978).
- Jagoutz, E. *et al.* The abundances of major, minor and trace elements in the Earth's mantle as derived from primitive ultramafic nodules. *Proc. Lunar Planet. Sci. Conf.* **2**, 2031–2050 (1979).
- Meisel, T., Walker, R. J. & Morgan, J. W. The osmium isotopic composition of the primitive upper mantle. *Nature* **383**, 517–520 (1996).
- Palme, H. & Beer, H. in *Astronomy and Astrophysics* (ed. Voigt, H. H.) 196–221 (Springer, Berlin, 1993).
- Jochum, K. P. Rhodium and other platinum-group elements in carbonaceous chondrites. *Geochim. Cosmochim. Acta* **60**, 3353–3357 (1996).
- Hertogen, J., Janssens, M. J., Takahashi, H., Morgan, J. W. & Anders, E. Enstatite chondrites; trace element clues to their origin. *Geochim. Cosmochim. Acta* **47**, 2241–2255 (1983).
- Kallemeyn, G. W. & Wasson, J. T. Compositions of enstatite (EH3, EH4.5 and EL6) chondrites; implications regarding their formation. *Geochim. Cosmochim. Acta* **50**, 2153–2164 (1986).
- Kong, P. & Ebihara, M. The origin and nebular history of the metal phase of ordinary chondrites. *Geochim. Cosmochim. Acta* **61**, 2317–2329 (1997).
- Cannat, M. *et al.* Site 920. *Proc. Ocean Drilling Progr. Part A: Init. Rep.* **153**, 45–119 (1995).
- Gillis, K. M. *et al.* Site 895. *Proc. Ocean Drilling Progr. Part A: Init. Rep.* **147**, 109–157 (1993).
- Snow, J. E. & Dick, H. J. B. Pervasive magnesium loss by marine weathering of peridotite. *Geochim. Cosmochim. Acta* **59**, 4219–4235 (1995).
- Snow, J. E. & Reisberg, L. Os isotopic systematics of the MORB mantle; results from altered abyssal peridotites. *Earth Planet. Sci. Lett.* **133**, 411–421 (1995).
- Peach, C. L., Mathez, E. A. & Keays, R. R. Sulfide melt–silicate melt distribution coefficients for noble metals and other chalcophile elements as deduced from MORB; implications for partial melting. *Geochim. Cosmochim. Acta* **54**, 3379–3389 (1990).
- Peach, C. L. & Mathez, E. A. Constraints on the formation of platinum-group element deposits in igneous rocks. *Econ. Geol. Bull. Soc. Econ. Geol.* **91**, 439–450 (1996).
- Reisberg, L. & Lorand, J. P. Longevity of sub-continental mantle lithosphere from osmium isotope systematics in orogenic peridotite massifs. *Nature* **376**, 159–162 (1995).
- O'Hanley, D. *Serpentinities: Records of Tectonic and Petrological History* 1–300 (Oxford Univ. Press, New York, 1996).
- Pattou, L., Lorand, J. P. & Gros, M. Non-chondritic platinum-group element ratios in the Earth's mantle. *Nature* **379**, 712–715 (1996).
- Schmidt, G., Palme, H., Kratz, K.-L. & Kurat, G. Are highly siderophile elements (PGE, Re and Au) fractionated in the upper mantle? New results on peridotites from Zabargad. *Chem. Geol.* (submitted).
- Anders, E. & Grevesse, N. Abundances of the elements: Meteoritic and solar. *Geochim. Cosmochim. Acta* **46**, 2363–2280 (1989).
- Kallemeyn, G. W. & Wasson, J. T. The compositional classification of chondrites-I: The carbonaceous chondrite groups. *Geochim. Cosmochim. Acta* **45**, 1217–1230 (1981).
- Wasson, J. T. & Kallemeyn, G. W. Compositions of chondrites. *Phil. Trans. R. Soc. Lond. A* **323**, 535–544 (1988).
- Walker, R. J., Morgan, J. W. & Horan, M. F. Osmium-187 enrichment in some plumes; evidence for core–mantle interaction? *Science* **269**, 819–822 (1995).
- Widom, E. & Shirey, S. Os isotope systematics in the azores: Implications for mantle plume sources. *Earth Planet. Sci. Lett.* **142**, 451–465 (1996).
- Walker, R. *et al.* Applications of the ^{190}Pt – ^{186}Os isotope system to geochemistry and cosmochemistry. *Geochim. Cosmochim. Acta* (in the press).
- Pernicka, E. & Wasson, J. T. Ru, Re, Os, Pt and Au in iron meteorites. *Geochim. Cosmochim. Acta* **51**, 1717–1726 (1987).
- Hoashi, M., Brooks, R. R. & Reeves, R. D. Palladium, platinum and ruthenium in iron meteorites and their taxonomic significance. *Chem. Geol.* **106**, 207–218 (1993).
- Ryan, D. E., Holzbecher, J. & Brooks, R. R. Rhodium and osmium in iron meteorites. *Chem. Geol.* **85**, 295–303 (1990).
- Schmidt, G. & Palme, H. in *EAG Workshop: The Origin and Fractionation of Highly Siderophile Elements in the Earth's Mantle* (eds Brüggemann, G., Lorand, J. P. & Palme, H.) 71–72 (Max-Planck Institut für Chemie, Mainz, Germany, 1997).
- Schmidt, G. Clues to the nature of the impacting bodies from platinum-group elements in borehole samples from the clearwater east crater (Canada) and the Boltysh impact crater (Ukraine). *Meteorit. Planet. Sci.* **32**, 761–767 (1997).
- Fryer, B. J. & Greenough, J. D. Evidence for mantle heterogeneity from platinum-group-element abundances in Indian Ocean basalts. *Can. J. Earth Sci.* **29**, 2329–2340 (1992).
- Morgan, J. P. & Morgan, W. J. Rare gas and isotopic evolution of the mantle. *Terra Abstr.* **9**, 47 (1997).
- Morgan, W. J. & Morgan, J. P. Two-stage melting of a multi-component mantle: A way to generate both OIB and MORB from 'whole mantle' convection. *Terra Abstr.* **9**, 57 (1997).
- Wasson, J. *Meteorites* 1–267 (Freeman, New York, 1985).
- Kallemeyn, G. W., Rubin, A. E., Wang, D. & Wasson, J. T. Ordinary chondrites; bulk compositions, classification, lithophile-element fractionations, and composition-petrographic type relationships. *Geochim. Cosmochim. Acta* **53**, 2747–2767 (1989).

Acknowledgements. We thank G. Brüggemann, A. W. Hofmann, K. P. Jochum and R. Walker for many useful discussions. Henry Dick and the Ocean Drilling Program kindly provided samples. L. Feld edited the typescript. J.E.S. acknowledges the support of an Alexander-von-Humboldt Fellowship during the course of this work. G.S. acknowledges financial support from the Deutsche Forschungsgemeinschaft to K.-L. Kratz and H. Palme.

Correspondence should be addressed to J.E.S. (e-mail: jesnow@geobar.mpgch-mainz.mpg.de).

Reactions between nitric oxide and haemoglobin under physiological conditions

Andrew J. Gow*† & Jonathan S. Stamler*

*Howard Hughes Medical Institute, Departments of Medicine and Cell Biology, Duke University Medical Center, Durham, North Carolina 27710, USA

†Institute for Environmental Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

The tenet of high-affinity nitric oxide (NO) binding to a haemoglobin (Hb) has shaped our view of haem proteins and of small diffusible signaling molecules. Specifically, NO binds rapidly to haem iron in Hb ($k \approx 10^7 \text{ M}^{-1} \text{ s}^{-1}$) (refs 1, 2) and once bound, the NO activity is largely irretrievable ($K_d \approx 10^{-5} \text{ s}^{-1}$) (refs 3–10); the binding is purportedly so tight as to be unaffected by O_2 or CO. However, these general principles do not consider the allosteric state of Hb or the nature of the allosteric effector, and they mostly derive from the functional behaviour of fully nitrosylated Hb, whereas Hb is only partially nitrosylated *in vivo*^{11–16}. Here we show that oxygen drives the conversion of nitrosylhaemoglobin in the 'tense' T (or partially nitrosylated, deoxy) structure to S-nitrosohaemoglobin in the 'relaxed' R (or ligand-bound, oxy) structure. In the absence of oxygen, nitroxyl anion (NO^-) is liberated in a reaction producing methaemoglobin. The yields of both S-nitrosohaemoglobin and methaemoglobin are dependent on the NO/Hb ratio. These newly discovered reactions elucidate mechanisms underlying NO function in the respiratory cycle, and provide insight into the aetiology of S-nitrosothiols, methaemoglobin and its related valency hybrids. Mechanistic re-examination of NO interactions with other haem proteins containing allosteric-site thiols may be warranted.

NO dissociation from the haems of Hb is much faster in the partially nitrosylated T structure than in the fully ligand-bound R structure^{3,4,17–19}. Moreover, in the T structure, the dissociation of NO is faster from the β -globin than from the α -globin chains^{16,18,19}. Thus, the equilibrium position in the T structure favours α -nitrosylhaemoglobin [$\text{Hb}(\alpha\text{FeNO}/\beta\text{Fe})$] over β -nitrosylhaemoglobin [$\text{Hb}(\alpha\text{Fe}/\beta\text{FeNO})$] (refs 16–19). We have shown previously that NO binds to cysteine 93 on the β -chain of Hb in the R structure and dissociates from the thiol in the T structure^{11,12}, and that the levels of nitrosyl Hb—which is most prevalent in deoxygenated venous blood^{11,14}—are inversely proportional to levels of S-nitrosohaemoglobin [$\text{Hb}(\beta\text{CysNO})$ or SNO-Hb], which is mainly found in oxygenated arterial blood^{11,12}. Taken together, these data on structural (that is T versus R) and chain (α versus β) non-equivalence in NO binding to Hb raise the possibility of an equilibrium between NO bound to haem in the T structure and to the β -chain thiol in the R structure.

To test for the existence of such an equilibrium, we examined the reactivity of nitrosyl Hb as a function of O_2 concentration. We chose NO-to-Hb ratios that are relevant to the high nanomolar to low micromolar concentrations of nitrosyl Hb found in physiological systems^{11–14}. Figure 1a shows the titration of partially nitrosylated deoxy Hb (NO:haem $\sim 1:70$) with air. The initial spectrum is a result of the combined absorbances of deoxy Hb (λ_{max} 430 nm) and nitrosyl Hb (λ_{max} 418 nm), with the former predominating. The gradual addition of air to this mixture produces the expected fall in absorbance at 430 nm (deoxy Hb) with a rise in absorbance at 415 nm (oxy Hb). However, the initial additions of air do not produce a true isosbestic point at around 421 nm, as would be expected if there was direct conversion of the one species to the other. This indicates that the concentration of at least one other species is changing. That said, a true isosbestic point is obtained

after the fifth addition of air. This strongly suggests that there is an alteration in the concentration of nitrosyl Hb [Hb(FeII)NO] mediated by the earlier additions of air.

We verified the change in the concentration of nitrosyl Hb by repeating this experiment, while following the absorbance difference compared with deoxy Hb. NO bound to the haems can be seen as a distinct peak in absorbance at 419 nm (Fig. 1b)⁵. Addition of air to this partially nitrosylated preparation results in a rise in absorbance at wavelengths below 420 nm (Fig. 1b), much like that seen in the absence of NO (Fig. 1c). At low levels of air, that is up to 30 μl , the 419 nm peak is preserved. But when more than 30 μl of air is added, the peak disappears (Fig. 1b). Thus NO is apparently released from the haems upon addition of sufficient air to promote the allosteric transition (from T to R) in Hb.

We reasoned that the loss of Hb(FeII)NO upon oxygenation results from the intramolecular transfer of the NO group to $\beta\text{-Cys93}$, forming SNO-Hb. We therefore measured the S-nitrosothiol (SNO) content of partially nitrosylated Hb, following oxygenation of solutions by vortexing in air. At relatively physiological ratios of NO to deoxy Hb (1 : 100 or less), a significant fraction of Hb(FeII)NO was converted to SNO-Hb(FeII)O_2 (Fig. 2). That is,

the yield of intramolecular NO group transfer from haem to thiol was ~70–85%. However, as the NO:Hb ratio was increased to 1 : 1, the yield of SNO decreased to zero; that is, NO became increasingly unavailable to form SNO. Once the NO concentration exceeded twice that of Hb, SNO was again observed; however, in this case, the yield reached a maximum of ~4% of the NO added (Fig. 2). Stated another way, NO group exchange between haem and thiol is very efficient when the O_2 -induced allosteric transition (T to R) in Hb drives the transfer of a physiological amount of NO. On the other hand, the efficiency of exchange drops off as the NO content of T-state Hb (that is, <2.7 NO/Hb (refs 2, 4, 17)) is increased. Once the initial concentration of NO is sufficient to maintain the R structure (that is, >2.7 NO/Hb (ref. 2)), SNO is again formed but with a lower percentage yield.

Our results support the idea of an equilibrium between NO bound to haems in the T structure and to thiols in the R structure, but some questions remain. In particular, why does the yield of SNO decrease when the initial NO : Hb ratio exceeds the physiological range? We looked for other NO reactions by examining the effect of sequential additions of NO to deoxy Hb (at 17.5 μM). The difference spectra (compared with deoxy Hb) in Fig. 3a show that physiolo-

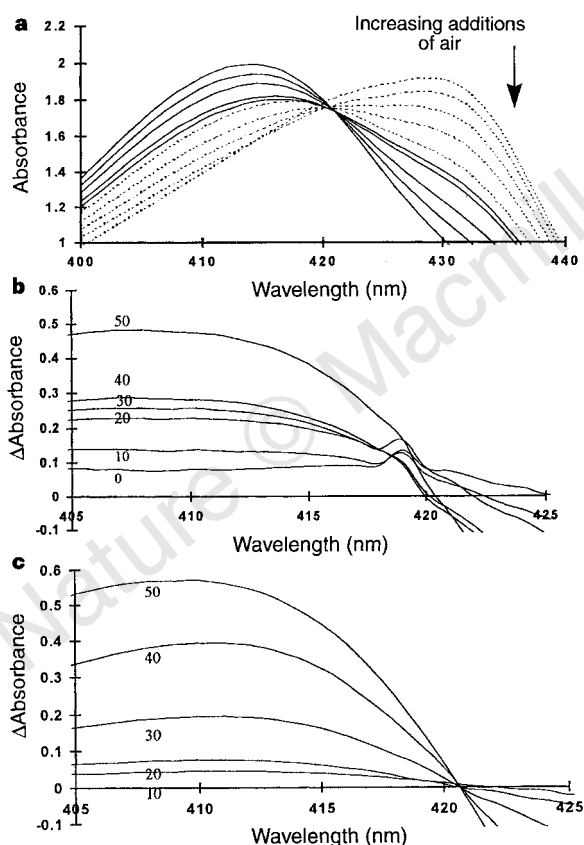
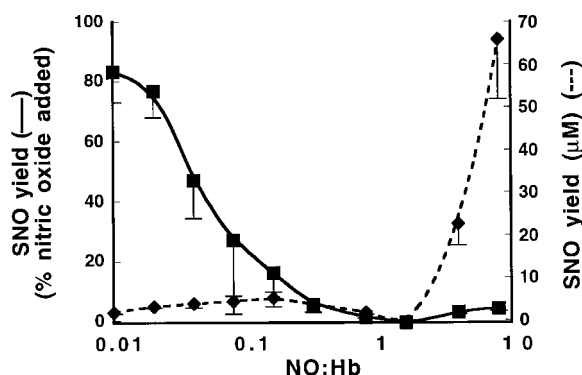


Figure 1 Oxygen titration of partially nitrosylated haemoglobin. **a**, Oxygen-driven dissociation of NO from haems of partially nitrosylated Hb. Nitric oxide was added to 17 μM deoxy Hb, such that the final NO : haem ratio was 1 : 68. The solution was then slowly aerated by sequential 50- μl injections of air. Initial additions of air failed to produce a true isosbestic point (dashed lines), indicating changes in the concentration of (at least) three species, deoxy, oxy and nitrosyl Hb. Later additions of air produced a true isosbestic point (solid lines), indicative of the conversion of deoxy Hb to oxy Hb. Therefore, there has been an apparent loss of nitrosyl Hb. **b**, Experimental outline in **a** was repeated using 20 μM Hb with a NO : haem ratio of 1 : 80 (NO/Hb $\approx$ 1 : 20). Aeration was achieved by sequential addition of 10 μl of air (numbers). Results are shown as difference spectra against deoxy Hb. The initial spectrum (0) is (the change in absorbance resulting from) partially nitrosylated Hb ($\Delta\lambda_{\text{max}}$ 419), so the absorbance is due to NO binding to the haems⁵. Early additions of air (up to 30 μl) result in mixtures of oxy and nitrosyl Hb. Later additions of air result in loss of nitrosyl Hb absorbance, as evidenced by disappearance of the peak at 419 nm. **c**, Difference spectra of air additions to 20 μM deoxy Hb. The absence of the nitrosyl Hb peak at 419 nm (as compared to **b**) is notable. The λ_{max} of the difference spectrum occurs at 410 nm, namely where the absorbance difference between deoxy Hb and oxy Hb is greatest, not at the λ_{max} of oxy Hb (415 nm). In addition, a true isosbestic point is found at ~420 nm where there is no absorbance (difference): absorbance rises at wavelengths below 420 nm and falls at wavelengths above 421 nm. Data in **a**, **b** and **c** are representative of three similar experiments.

Figure 2 Auto S-nitrosylation of haemoglobin upon oxygenation: NO group transfer from haem iron to cysteine as a function of initial NO/Hb ratio. Deoxy Hb was nitrosylated to various degrees and then immediately oxygenated by vortexing in air. The amount of SNO that forms is presented both as a percentage of the amount of NO initially added to deoxy Hb (200 μM) (solid line) and as an absolute amount (μM ; dashed line). Yields of SNO are up to 12% lower if calculated based on direct nitrosyl haem measurements¹¹, that is, NO group transfer from haem to thiol may be as much as 12% less efficient. This difference may reflect some higher oxide of nitrogen (NO_x) contamination of NO solutions and/or an additional route of Hb S-nitrosylation. Data are the mean $\pm$ s.e.m. of 4–27 experiments at each point.



gical concentrations of NO ($<1 \mu\text{M}$) produce an increase in absorbance below 420 nm, with a peak in the region of 408–412 nm. The spectrum is indicative of partially nitrosylated deoxy Hb. Surprisingly, addition of more NO produces a shift in the spectrum towards 404 nm, indicating oxidation to methaemoglobin (met Hb)⁵. Peak absorption at this wavelength occurs when the

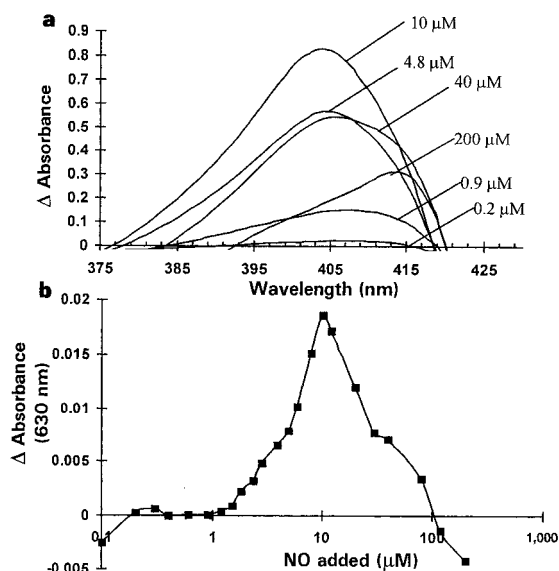


Figure 3 Titration of deoxyhaemoglobin with nitric oxide: redox reactions of NO bound to the haem iron. **a**, Haem oxidation and reduction by sequential addition of NO under anaerobic conditions. Addition of low concentrations of NO (0.2–0.9 μM) to deoxy Hb (17.5 μM) produce partially nitrosylated Hb (seen as a peak in the difference spectrum at $\sim 414 \text{ nm}$). (In difference spectra against deoxy Hb, the maximal difference peak does not necessarily correspond with the λ_{max} (nitrosyl Hb, 418 nm) of the species being formed.) Addition of intermediate concentrations of NO (4.8–10 μM) produce met Hb (peak at $\sim 404 \text{ nm}$). Addition of higher concentrations of NO (40–200 μM) then reduce met Hb and bind the haems to form nitrosyl Hb (peak shifts back to $\sim 415 \text{ nm}$). Thus, charge transfer reactions of NO at the haem iron produce met Hb and NO^- (equations (3, 4)), after which excess NO reduces met Hb through the intermediacy of NO^- (equations (1, 2)). Data are presented as difference spectra in the Soret region against doxy Hb. **b**, Met Hb production under anaerobic conditions as a function of NO concentration. Repeat of the experiment shown in **a**, while following the absorbance change at 630 nm (met Hb). Data in **a** and **b** are of representative of four similar experiments.

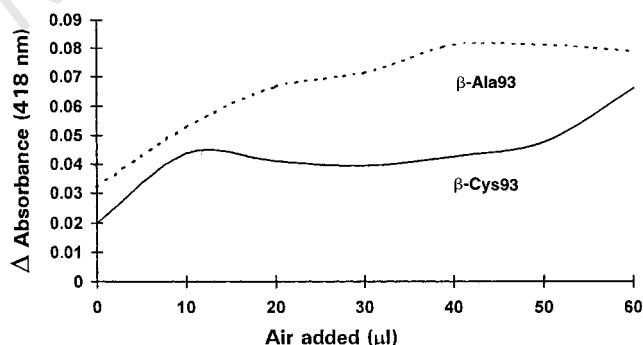
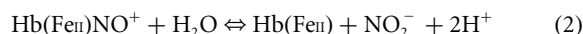
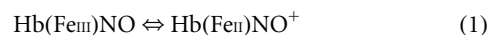


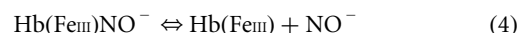
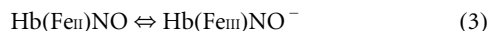
Figure 4 Oxygen-driven decomposition of nitrosyl haemoglobin is dependent on β -chain cysteine 93. The O_2 titration of partially nitrosylated Hb, as for Fig. 1, was repeated using recombinant haemoglobins having either a cysteine (wild-type; solid line) or an alanine (mutant; dashed line) at position $\beta 93$. Oxygen-mediated NO dissociation from the haem iron occurred in the wildtype but not in the mutant. The absorbance at 418 nm minus the absorbance of deoxy Hb is shown. The result suggests apparent coupling of the O_2 -driven allosteric transition with exchange of the NO group between haem and thiol. Recombinant haemoglobins were provided by C. Fronticelli. Data are representative of five similar experiments.

NO:Hb ratio is $\sim 1:2$ (50%). Further addition of NO results in a rightward shift of the peak to 414 nm, indicating the formation of nitrosyl Hb⁵. The accompanying fall in the peak height is consistent with nitrosyl Hb having a lower extinction coefficient than met Hb⁵. These results indicate that nitrosyl Hb is formed at low ratios of NO:Hb ($<1:20$), met Hb is formed in the mid-range ($1:20 < \text{NO}:\text{Hb} < 1:2$), and nitrosyl Hb is formed at high ratios ($\text{NO}:\text{Hb} > 1:2$).

To confirm the formation of met Hb from mid-range additions of NO to deoxy Hb (17.5 μM), we monitored the change in absorbance at 630 nm as a function of NO added⁵ (Fig. 3b). In agreement with our previous result, met Hb did not form when less than 1 μM NO was added and was maximal when the NO to deoxy Hb ratio was 57% (an amount of NO that is 'insufficient' to effect the allosteric transition²). That is, met Hb does not form at ratios of NO to Hb that are conducive to S-nitrosothiol formation ($<1:25$), but does so at higher ratios within the T structure (up to $1:2 \text{ NO/Hb}$); the amount of met Hb produced was ~ 30 –45% of the NO added, arguing against trace O_2 playing a role. The subsequent fall in met Hb concentration with addition of $>10 \mu\text{M}$ NO (that is, $>1:2 \text{ NO/Hb}$) is the result of met Hb reduction by NO (ref. 20) (equations (1, 2)). Nitrite is a by-product of this reaction (equation (2)) and may contaminate NO solutions, but control titrations excluded any major involvement in met Hb formation (data not shown).



The mechanism of met Hb generation is best rationalized by charge transfer reactions at the haem; heterolytic cleavage of the Fe(II)NO bond would then produce Fe(III) and nitroxyl anion (NO^-) (ref. 21) (equations (3, 4))



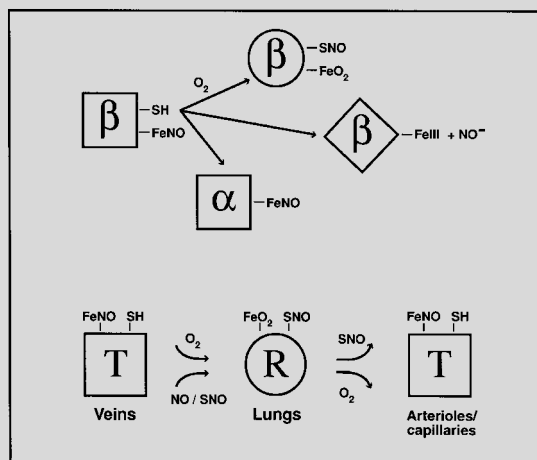
The generation of NO^- was verified by measurement of both N_2O (the product of NO^- protonation, dimerization and dehydration) and hydroxylamine (the product of singlet NO^- reaction with thiol)^{21,22}. Mass spectrometry showed a peak of relative molecular mass 44 (corresponding to N_2O), which was dependent on the concentration of NO (consistent with Fig. 3b), when the gas phase above the reaction was analysed²². Hydroxylamine formation was assayed by indoxine formation from 8-hydroxyquinoline following addition of 10 mM *N*-acetylcysteine (or cysteine) to reaction mixtures²². Hydroxylamine was only detectable after additions of $>5 \mu\text{M}$ NO ($\text{NO}:\text{Hb} > 1:3.5$). It has been shown that binding of 2.7 NO molecules per Hb tetramer is required to elicit the conformational change from T to R (ref. 2), so Fe(III) haem and NO^- probably derive from a reactive Fe(II)NO bond in the T state.

To verify the role of β -Cys93 in NO dissociation, we examined the functional behaviour of a mutant HbA in which the cysteine at position β -93 was replaced by alanine (β -Ala93); the recombinant wild-type Hb (β -Cys93) was used as the control (Fig. 4). Both haemoglobins form a nitrosyl complex on NO addition (λ_{max} 418 nm). At a NO:haem ratio of $\sim 1:72$ (T structure), early additions of air produce a similar rise in absorbance at 418 nm for both haemoglobins (which is caused by O_2 binding to the deoxy haems⁵). The later additions of air, however, result in a slight fall and then plateau in absorbance for β -Cys93, during which time the absorbance of β -Ala93 continues to rise. The fall/plateau in the wild-type Hb is indicative of NO dissociation from the haem. On the other hand, the thiol appears to play a lesser role in NO dissociation when the haemoglobins are fully nitrosylated to begin with (that is, in the R structure). In this case, the addition of air produces only a very small drop of $\sim 5\%$ in the peak at 418 nm (not shown), consistent with the amounts converted to SNO (Fig. 2). A similar amount of β -chain nitrosyl haem is slowly

formed after SNO-Hb is synthesized by transnitrosation (D. J. Singel, B. Luchsinger and J.S.S., manuscript in preparation). S-nitrosothiol was not detectable after oxygenation of T-structure nitrosyl β -Ala93Hb. These data indicate that (1) the stability of nitrosyl haem is regulated by the thiol at position β 93; (2) β -chain nitrosyl haems of fully nitrosylated Hb in the R structure are less

Box 1 Nitric oxide reactions with haemoglobin in the respiratory cycle

The upper panel in the figure shows alternative reactions proposed for β -chain nitrosyl haems in the T structure. (1) Transfer of NO to the α -chain haem irons (this is likely to occur mainly in the microcirculation and venous system); (2) charge-transfer reaction at the haem iron to produce methaemoglobin and nitroxyl anion (this is more likely to occur in the microcirculation and venous system when NO synthesis is high); and (3) NO group exchange with β -Cys93 mediated by the oxygen-driven allosteric transition to the R structure, forming SNO-Hb(FeII)O₂ (this



seems to occur in the lung^{11,12}, but may also happen in oxygenated arterial blood). Squares, circle and diamond represent haems in the T structure, R structure and met state, respectively. The lower panel shows a model of NO binding to haems and thiols of Hb in the circulation. Partially nitrosylated venous blood enters the lungs in the deoxy or T structure (square) (presumably maintained by α -chain binding of NO to haems and CO₂ to amines). Specifically, NO can be detected in large veins as α (5- and 6-coordinate) nitrosyl haem^{13,14,16,24}, some of which is found in the fraction of blood that is fully deoxygenated. In the lungs, partially nitrosylated (carbamino) Hb is exposed to more NO and O₂ tensions (pO_2) that appear to couple the allosteric transition with NO group exchange from haems to β -chain thiols^{11,12}. Oxygen serves both to position the β thiol close to the β haem^{12,25} and thermodynamically to drive the redox-mediated formation of SNO. Accordingly, blood entering the arterial circuit contains SNO-oxy Hb, that is, Hb in the R structure (circle) with NO attached to β -Cys93 and O₂ to the haems. Low-molecular-mass SNOs present in the lung and arterial blood will further support SNO-Hb formation by transnitrosation of R-structured molecules¹¹. Blood moving into resistance vessels that control blood pressure and blood flow to tissues is then exposed to low pO_2 which promotes the T structure in SNO-Hb and effects NO group release. Some NO will exchange with low-relative-mass thiols to dilate blood vessels^{11,12} and some will be autocaptured at the haems ($\beta \rightarrow \alpha$)¹¹ (just as some endothelial-derived NO is inevitably sequestered by the haems^{11,12}); NO oxidation of haem irons (met formation) will also enhance the vasodilator function of SNO-Hb¹¹. As O₂ delivery is a function of blood flow, the R $\rightarrow$ T transition in Hb (and perhaps met Hb formation) is designed to maximize oxygen delivery. Hb can further bind NO at the haem irons as it progresses through the venous system; the more NO that binds, the greater the propensity to form met Hb and haemoglobin X [Hb(α FeII NO)(β FeIII)]. It is not known whether the endogenous level of ~0.1% nitrosyl Hb¹², which should promote the T structure in Hb, is sufficient to enhance O₂ delivery, but higher levels found in endotoxic shock¹⁶ may do so.

reactive than β -chain nitrosyl haems in T structure, consistent with earlier reports^{16,18}; and (3) NO group exchange occurs between the haems and thiols. Indeed, Hb effectively catalyses SNO formation under physiological conditions, that is, it converts 'substrate' NO (in the presence of 'cosubstrate' O₂) into 'product' SNO, whose release from the protein is facilitated by glutathione^{11,12}.

We offer an expanded model of NO-Hb interactions that integrates these new data. NO rapidly associates with either the α or β haems of T-state Hb, but dissociates much faster from the β -chains^{5,6,16,18,19}. This manifests in rapid redistribution of β nitrosyls to the α haems^{16,18,19}. As more NO binds to Hb in the T structure, the β nitrosyl haems [Fe(II)NO] cleave heterolytically to produce NO⁻ and Fe(III). (Six-coordinate α nitrosyl haems may cleave heterolytically as well, but more stable 5-coordinate α nitrosyls are unlikely to do so). In other words, heterolytic cleavage is increasingly favoured over homolytic NO dissociation as the equilibrium position for NO between α/β haems is reached (Box 1, top figure). These alternative reactions provide new insight into the aetiology of met Hb and related valency hybrids such as haemoglobin X (Hb(α FeII NO)(β FeIII))^{13-15,23} in blood. They also explain why NO is bound selectively to α haems in the venous circulation^{11,12,16,24}.

The allosteric transition in Hb produces a change in the position and reactivity of β -Cys93 (refs 12, 25, 26). The thiol, which is unreactive and distanced from the β haem in the T structure, is activated and brought into close proximity in the R structure¹². NO can then transfer from the haem to the thiol. The remaining problem in the formation of nitrosothiols is the need for an electron acceptor^{21,27}. Oxygen can fulfil this requirement (whether the reaction involves the intermediate RSNO⁻ (ref. 27) or the formal transfer of NO⁺ (ref. 21)): it is uniquely capable, among small diffusible ligands, of both promoting the allosteric transition and supporting the reaction yielding SNO (in which case the product is superoxide). In this view, oxygen dictates the position of an equilibrium that exists between NO bound to β haem and β thiol²⁸ (Box 1, top figure). This model rationalizes both the inability to detect NO bound to β haems in arterial blood, and the predominance of arterial SNO^{11-14,16}.

The chemistry of partially nitrosylated (T structure) Hb observed *in vitro*—namely, O₂-induced ligand exchange from haem to thiol coupled with the allosteric transition—has parallels *in vivo*. In central venous blood, NO resides predominantly on T-state haems^{11,14}; after oxygenation in the lung, the NO group is found partly on R-state thiols^{11,12} (Box 1, bottom figure). Inasmuch as NO bound to Hb remains constant during this cycling^{11,12}, NO groups appear to migrate from haems to thiols. SNO-Hb then moves through the systemic circulation until it encounters the physiological O₂ gradient in resistance vessels that promotes the T structure in Hb and effects (S)NO release^{11,12} (Box 1, bottom figure). Some NO group transfer to glutathione and smooth muscle occurs to relax blood vessels and thereby promotes O₂ delivery^{11,12}. The other NO groups are evidently transferred back to the haems^{11,12,16,24}. In this way, the allosteric transition (R to T) regulating blood flow would be reinforced in SNO-carrying molecules (the negative *trans* effect^{6,17} leads to cooperativity of SNO dissociation). The migration of NO back to haems could also function to maintain the T structure in Hb molecules as they become exposed to relatively high O₂ tensions in the large veins (negative cooperativity of O₂ association). Haemoglobin might otherwise undergo untimely allosteric transitions and/or reactions that would deplete the NO pool. NO may exchange between haems and cysteines in other proteins^{21,29,30}. Such reactions might illuminate new aspects of metalloprotein biochemistry and cellular physiology. □

Methods

Preparation of deoxy Hb and nitrosyl Hb. Solutions of Hb in 0.1 M phosphate, 100 μ M diethylenetriaminepentaacetic acid (DTPA), pH 7.4, in a 2-ml sealed vial were deoxygenated by bubbling in argon for 45 min; deoxygenation was

verified spectrophotometrically. Nitric oxide was added from a stock solution, prepared under argon and purified across NaOH and a cold trap, by injection through a gas-tight Hamilton syringe with a Teflon seal. The concentration of nitric oxide stock solutions varied between 1.5 and 1.9 mM, as determined by a NO-specific electrode and oxidation of haemoglobin.

Titration of partially nitrosylated Hb with O₂. Oxygenation of Hb was achieved by sequentially injecting either 10 or 50 μ l of air into a 2-ml sealed vial with a gas-tight Hamilton syringe.

S-nitrosylation of Hb. Varying amounts of NO were added to deoxy Hb (200 μ M) for a final Hb concentration of 100–200 μ M (higher Hb concentrations pose technical problems in this experimental design) and vortexed under argon for a few seconds. Vials were then unsealed and vortexed in air until well oxygenated. The SNO-Hb yield was measured as described¹¹, with the exception that HgCl₂ was used at 6–7.5-fold the Hb concentration.

Titration of deoxy Hb with NO. 1 ml deoxy Hb (17.5 μ M) was prepared in a sealed quartz cuvette. NO was added sequentially from a stock solution by injection through a gas-tight Hamilton syringe with Teflon seal, and vortexed. Spectra were recorded after each addition. Difference spectra were calculated by subtracting the initial deoxy-Hb spectrum.

Received 5 August; accepted 29 September 1997.

- Gibson, Q. H. & Rougton, F. J. W. The kinetics and equilibria of the reactions of nitric oxide with sheep haemoglobin. *J. Physiol.* **136**, 507–526 (1957).
- Cassidy, R. & Gibson, Q. H. Conformation, co-operativity and ligand binding in human hemoglobin. *J. Mol. Biol.* **91**, 301–313 (1975).
- Sharma, V. S. & Ranney, H. M. The dissociation of NO from nitrosylhemoglobin. *J. Biol. Chem.* **253**, 6467–6472 (1978).
- Moore, E. G. & Gibson, Q. H. Cooperativity in the dissociation of nitric oxide from hemoglobin. *J. Biol. Chem.* **251**, 2788–2794 (1976).
- Kharitonov, V. G., Bonaventura, J. & Sharma, V. S. in *Methods in Nitric Oxide Research* (eds Feelisch, M. & Stamler, J. S.) 39–47 (Wiley, London, 1996).
- Traylor, T. G. & Sharma, V. S. Why NO? *Biochemistry* **31**, 2847–2849 (1992).
- Lancaster, J. R. Simulation of the diffusion and reaction of endogenously produced nitric oxide. *Proc. Natl Acad. Sci. USA* **91**, 8137–8141 (1994).
- Motterlini, R., Vandegriff, K. D. & Winslow, R. M. Hemoglobin–nitric oxide interaction and its implications. *Transfus. Med. Rev.* **X**, 77–84 (1996).
- Marletta, M. A., Tayeh, M. A. & Hevel, J. M. Unraveling the biological significance of nitric oxide. *Biofactors* **2**, 219–225 (1990).
- Antonini, E. & Brunori, M. in *Hemoglobin and Myoglobin in their Reactions with Ligands* **32**, 272–276 (Elsevier, New York, 1971).
- Jia, L., Bonaventura, C., Bonaventura, J. & Stamler, J. S. S-nitrosohemoglobin: a dynamic activity of blood involved in vascular control. *Nature* **380**, 221–226 (1996).
- Stamler, J. S. *et al.* Blood flow regulation by S-nitrosohemoglobin is controlled by the physiological oxygen gradient. *Science* **276**, 2034–2037 (1997).
- Westenberger, U. *et al.* Formation of free radicals and nitric oxide derivative of hemoglobin in rats during septic shock. *Free Rad. Res. Commun.* **11**, 167–178 (1990).
- Hall, D. M., Buettner, G. R., Mathes, R. D. & Gisolfi, C. V. Hyperthermia stimulates nitric oxide formation: electron paramagnetic resonance detection of NO-heme in blood. *J. Appl. Physiol.* **77**, 548–553 (1994).
- Kruszyna, R., Kruszyna, H., Smith, R. P. & Wilcox, D. E. Generation of valency hybrids and nitrosylated species of hemoglobin in mice by nitric oxide vasodilators. *Toxicol. Appl. Pharmacol.* **94**, 458–465 (1988).
- Kosaka, H. *et al.* ESR spectral transition by arteriovenous cycle in nitric oxide hemoglobin of cytokine-treated rats. *Am. J. Physiol.* **266**, 1400–1405 (1994).
- Perutz, M. F. in *Molecular Basis of Blood Diseases* (ed. Stamatoyannopoulos, G.) 127–178 (Saunders, Philadelphia, 1987).
- Taketa, F., Antholine, W. E. & Chen, J. Y. Chain nonequivalence in binding of nitric oxide to hemoglobin. *J. Biol. Chem.* **253**, 5448–5451 (1978).
- Henry, Y. *et al.* EPR characterization of molecular targets for NO in mammalian cells and organelles. *FASEB J.* **7**, 1124–1134 (1993).
- Addison, A. W. & Stephanos, J. J. Nitrosyliron (III) hemoglobin: autoreduction and spectroscopy. *Biochemistry* **25**, 4104–4113 (1986).
- Stamler, J. S., Singel, D. J., Loscalzo, J. Biochemistry of nitric oxide and its redox activated forms. *Science* **258**, 1898–1902 (1992).
- Arnelle, D. & Stamler, J. S. NO⁺, NO[•], and NO[−] donation by S-nitrosothiols: Implications for regulation of physiological functions by S-nitrosylation and acceleration of disulfide formation. *Archiv. Biochem. Biophys.* **318**, 279–285 (1995).
- Kruszyna, R., Kruszyna, H., Smith, R. P., Thron, R. D. & Wilcox, D. E. Nitrite conversion to nitric oxide in red cells and its stabilization as a nitrosylated valency hybrid of hemoglobin. *J. Pharmacol. Exp. Ther.* **241**, 307–313 (1987).
- Kagan, V. E., Day, B. W., Elsayed, N. V. & Gorbunov, N. V. Dynamics of nitrosylated haemoglobin in blood. *Nature* **383**, 30–31 (1996).
- Riggs, A. & Wolbach, R. A. Sulfhydryl groups and the structure of hemoglobin. *J. Gen. Physiol.* **39**, 585–605 (1962).
- Riggs, A. F. The binding of N-ethylmaleimide by human hemoglobin and its effect upon the oxygen equilibrium. *J. Biol. Chem.* **236**, 1948–1954 (1961).
- Gow, A., Buerk, D. D. & Ischiropoulos, H. A novel reaction mechanism for the formation of S-nitrosothiol in vivo. *J. Biol. Chem.* **272**, 2841–2845 (1997).
- Perutz, M. F. Blood: Taking the pressure off. *Nature* **380**, 205–206 (1996).
- Stamler, J. S., Toone, E. J., Lipton, S. A. & Sucher, N. J. (S)NO Signals: Translocation, regulation, and a consensus motif. *Neuron* **18**, 691–696 (1997).
- Stamler, J. S. Redox signalling: Nitrosylation and related target interactions of nitric oxide. *Cell* **78**, 931–936 (1994).

Acknowledgements. We thank J. Bonaventura, D. J. Singel, H. Ischiropoulos and I. Fridovich for discussion and T. McMahon for help with measurements. J.S.S. is the recipient of grants from the NHLBI; A.J.G. is supported by a National Research Award from the NHLBI.

Correspondence and requests for materials should be addressed to J.S.S.

Acetylcholine receptors containing the $\beta 2$ subunit are involved in the reinforcing properties of nicotine

Marina R. Picciotto^{*†‡}, Michele Zoli^{*†‡}, Roberto Rimondini[§], Clément Léna^{*}, Lisa M. Marubio^{*}, Emilio Merlo Pich^{||}, Kjell Fuxe[§] & Jean-Pierre Changeux^{*}

^{*} CNRS UA D1284, Neurobiologie Moléculaire, Institut Pasteur, 28 rue du Dr Roux, 75724 Paris Cédex 15, France

[§] Department of Histology, Karolinska Institute, Box 60400, S-10401 Stockholm, Sweden

^{||} Geneva Biomedical Research Institute, Glaxo-Wellcome R&D, CH-1228 Geneva, Switzerland

[†] These authors contributed equally to this work.

Release of the neurotransmitter dopamine in the mesolimbic system of the brain mediates the reinforcing properties of several drugs of abuse, including nicotine¹. Here we investigate the contribution of the high-affinity neuronal nicotinic acetylcholine receptor² to the effects of nicotine on the mesolimbic dopamine system in mice lacking the $\beta 2$ subunit of this receptor³. We found that nicotine stimulates dopamine release in the ventral striatum of wild-type mice but not in the ventral striatum of $\beta 2$ -mutant mice. Using patch-clamp recording, we show that mesencephalic dopaminergic neurons from mice without the $\beta 2$ subunit no longer respond to nicotine, and that self-administration of nicotine is attenuated in these mutant mice. Our results strongly support the idea that the $\beta 2$ -containing neuronal nicotinic acetylcholine receptor is involved in mediating the reinforcing properties of nicotine.

The mesostriatal dopaminergic system influences many important behaviours, such as locomotor activity and the reinforcing action of many drugs of abuse, including ethanol, cocaine and amphetamine⁴. Both the locomotor-stimulating action and the reinforcing properties of nicotine are thought to be mediated through increasing dopamine release in the mesolimbic dopamine system⁵. Nicotine administered systemically increases extracellular levels of dopamine in the dorsal and ventral (nucleus accumbens) striatum⁶ by acting on the nicotinic acetylcholine receptors (nAChRs) located on dopamine cell bodies and/or nerve terminals⁷. Accordingly, lesions of the mesolimbic dopaminergic neurons attenuate the locomotor-stimulating properties of nicotine⁸ and nicotine self-administration in rats⁹. The ten subunits of the neuronal nAChR identified so far combine to form a variety of receptor subtypes that are sensitive to nicotine, and at least six of these subunits ($\alpha 3$, $\alpha 4$, $\alpha 5$, $\alpha 6$, $\beta 2$ and $\beta 3$) are expressed in mesencephalic dopaminergic neurons¹⁰. We therefore investigated which subtypes of the nAChR regulate dopamine levels in response to systemically administered nicotine and so contribute to the properties of nicotine that are likely to be involved in tobacco consumption.

We used 'knockout' mice that lacked the $\beta 2$ subunit of the nAChR³ in order to investigate nicotine-elicited dopamine release and related behaviours. We have already shown that these mutant mice have normal levels of the messenger RNAs encoding the $\alpha 2$ – $\alpha 7$ and $\beta 3$, $\beta 4$ subunits of the nAChR, but that they lack high-affinity binding of nicotine in the brain³. The mutant animals are indistinguishable from their wild-type siblings in a cage, but they lose their sensitivity to nicotine, as judged by their performance in a

[†] Present addresses: Department of Psychiatry, Yale University School of Medicine, 34 Park Street, New Haven, Connecticut 06508, USA (M.R.P.); Department of Biomedical Sciences, Section of Physiology, University of Modena, 41100 Modena, Italy (M.Z.).

passive avoidance-learning task, and apparently acquire this behaviour more easily. In a first series of experiments, we used *in vivo* microdialysis to detect dopamine and its metabolites in $\beta 2$ -mutant mice and their wild-type siblings and to determine whether nicotine still elicits an increase in extracellular dopamine in the mutant animals. A 2-mm dialysis probe was used which spanned the ventral and part of the dorsal mouse striatum, and sampled both limbic and motor regions. Nicotine (free base) was administered intraperitoneally (i.p.) at 3 doses (0.030, 0.125 and 0.500 mg kg⁻¹) to anaesthetized mice and the dopamine concentration was measured in the dialysate for 120 min (Fig. 1). Nicotine doses were chosen after extensive testing of DBA/2 mice, C57/B16 mice and an F₁ cross

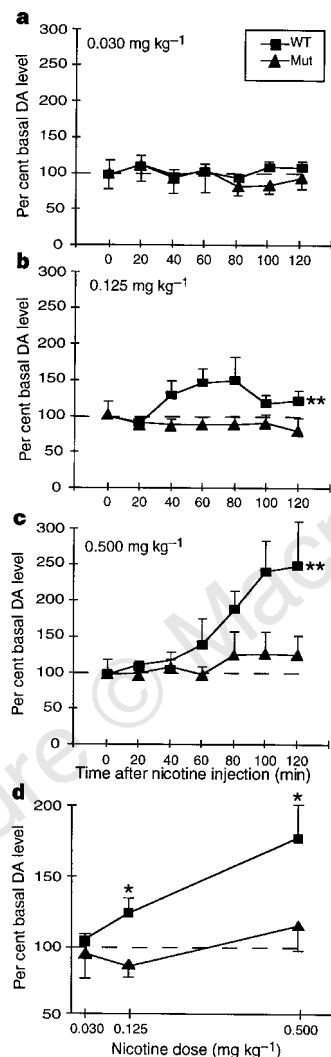


Figure 1 Nicotine-elicited dopamine (DA) release in $\beta 2$ -mutant mice and their wild-type siblings. **a–c**, Mice were injected i.p. with a solution of 0.030, 0.125 or 0.500 mg kg⁻¹ of nicotine (free base), and striatal extracellular DA levels were measured by microdialysis. Male mice aged 3–5 months were used. Results are reported as a percentage of the mean of three basal samples immediately before treatment. Basal dopamine levels in perfusate did not differ between the two groups: for mutant, 2.94 ± 0.38 nM; for wild type, 3.07 ± 0.49 nM. Means $\pm$ s.e.m. are shown (n , 6–8 animals per group). Statistical analysis was carried out by a repeated-measures, two-way analysis of variance (ANOVA); ** $P < 0.01$, for wild-type versus mutant for both 0.125 and 0.500 mg kg⁻¹ doses. **d**, Dose-response curve of the effect of nicotine on extracellular dopamine levels in $\beta 2$ -mutant mice and their wild-type siblings after i.p. injection with 0.030, 0.125 or 0.500 mg kg⁻¹ nicotine. Average per cent changes compared to basal in the 6 samples collected after nicotine injection are shown. Statistics according to one-way ANOVA; * $P < 0.05$, mutant compared with wild type. WT, +/+ mice; Mut, homozygous $\beta 2$ -mutant mice.

of these two strains, the strains included in the genetic background of the $\beta 2$ -mutant mice and their wild-type siblings³. The maximally effective dose of nicotine was 0.500 mg kg⁻¹ for dopamine release in C57B16 mice, whereas 0.030 mg kg⁻¹ of nicotine was ineffective in all strains. These doses are in the range used in studies of behavioural pharmacology and dopamine microdialysis in rodents^{6,11,12}. In $\beta 2$ ^{+/+} mice, 0.500 mg kg⁻¹ and 0.125 mg kg⁻¹ nicotine caused a peak increase of ~ 150 and 50%, respectively, in dopamine in the

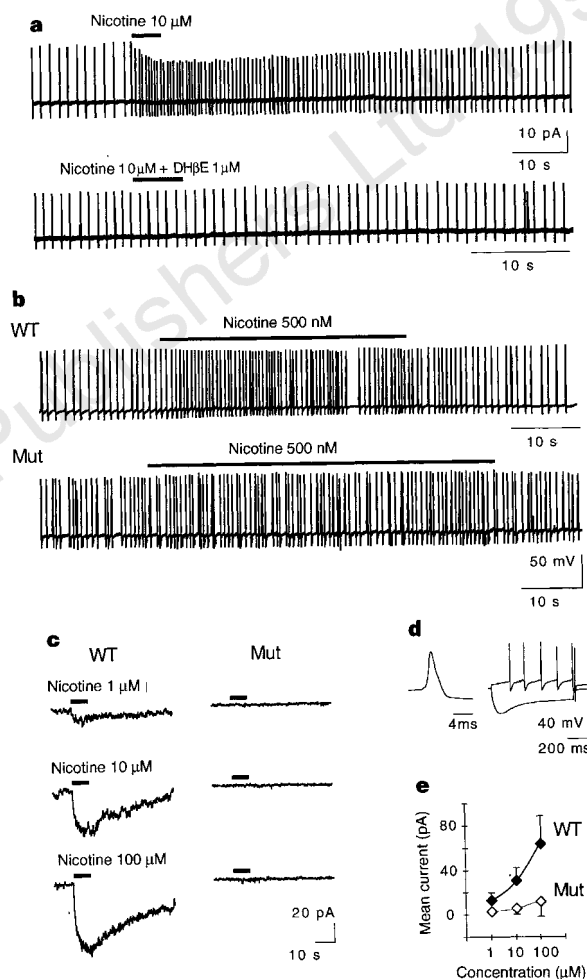


Figure 2 Patch-clamp recording of nicotine-evoked currents in substantia nigra (SN) pars compacta and ventral tegmental area (VTA) of wild-type and $\beta 2$ ^{-/-} mice. **a**, 10 μ M nicotine greatly increases the frequency of discharge of SN neurons recorded in the cell-attached configuration in neurons from wild-type (WT) mice. The nicotinic antagonist DH β E (1 μ M) reversibly blocks the action of nicotine in neurons from +/+ animals. **b**, 500 nM nicotine moderately increases the firing rate of dopaminergic neurons from wild-type mice in the whole-cell configuration under current clamp conditions ($P \leq 0.001$). No increase in firing rate is seen in dopaminergic neurons from $\beta 2$ ^{-/-} mice. Statistical significance of the change in frequency in wild-type and mutant mice was tested with the matched *t*-test and the Wilcoxon test on the matched pairs of frequency value before and during nicotine application. **c**, Nicotine (1–100 μ M) elicits an inward current in SN and VTA neurons recorded in whole-cell configuration in wild-type but not in homozygous mutant mice. The slight increase in baseline noise after 100 μ M agonist application in mutant mice indicates that another type of nAChR may be present on these neurons. **d**, Broad spikes and characteristic response to hyperpolarizing/depolarizing current pulses of dopaminergic neurons. The left trace is a single action potential at expanded timescale; right trace corresponds to current injection (–250 pA, +150 pA) into the cell. Vertical scale is the same for the left and right traces. **e**, Dose-response curve for nicotine in wild-type and mutant mice. Values correspond to mean $\pm$ s.d. of currents recorded at –70 mV. Each point corresponds to the average of 3 to 10 measures.

perfusate (Fig. 1a–d). In contrast to the results obtained with wild-type mice, dopamine levels were not significantly changed in the striatum of $\beta 2$ -mutant mice in response to any dose of nicotine (Fig. 1a–d). These results show that the $\beta 2$ subunit is a necessary component of the receptor(s) that are primarily responsible for the pharmacological effect of systemic nicotine on striatal dopamine release.

nAChRs located on mesencephalic dopaminergic cell bodies are necessary for dopamine release elicited by systemic nicotine administration¹³, and release of dopamine in the nucleus accumbens is likely to be involved in the reinforcing properties of nicotine⁶. We therefore investigated the effect of nicotine on the discharge frequency of dopaminergic neurons in thin slices of the substantia nigra (SN) and ventral tegmental area (VTA) in both cell-attached (Fig. 2a) and whole-cell (Fig. 2b) configurations to see whether there was a cellular response to nicotine in dopaminergic neurons from $\beta 2^{-/-}$ mice. Nicotine at 10 μ M increased the frequency of discharge of identified dopaminergic neurons (3/3 cells) and this effect was blocked by 1 μ M DH β E, a competitive nicotinic antagonist (Fig. 2a). We also tested the effect of 500 nM nicotine, which is close to the concentration of nicotine in the blood of human smokers (~500 nM nicotine from one puff¹⁴) and in rats injected i.p. with 0.1 mg kg⁻¹ nicotine¹⁵. In wild-type mice, 500 nM nicotine increased the discharge frequency in 10 out of 15 cells recorded (Fig. 2b). In contrast, no response was seen in 14 out of 15 dopaminergic neurons from $\beta 2^{-/-}$ mice (Fig. 2b). The average increase in discharge frequency following the application of 500 nM nicotine was 2.5 ± 0.5 ($n = 15$) in wild-type mice and 1.1 ± 0.3 ($n = 15$) in $\beta 2^{-/-}$ mice. This indicates that most, if not all, of the responses were due to the activation of $\beta 2$ -subunit-containing nAChRs. All cells recorded showed the characteristic broad spikes and response to hyperpolarizing and depolarizing current pulses characteristic of dopaminergic neurons¹⁶ (Fig. 2d). The lack of response to nicotine in these neurons from $\beta 2^{-/-}$ mice was confirmed in voltage-clamp experiments in which neurons in the SN and VTA of wild-type mice responded consistently to 1, 10 and 100 μ M nicotine with a small but detectable inward current ($n = 46$) (Fig. 2c, e) which was also blocked by 1 μ M DH β E (not shown). The agonist order of the response was compatible with that of $\alpha 4/\beta 2$ -containing nicotinic receptors *in vitro* (nicotine > cytosine) (results not shown)¹⁷. In 19 of 20 neurons tested from $\beta 2^{-/-}$ mice, the response of SN and VTA neurons to nicotine was abolished (Fig. 2c, e). Nicotine caused a small inward current in 1 of 20 neurons, however, and 100 μ M nicotine and cytosine could raise baseline noise in a few cells, indicating that nAChRs lacking the $\beta 2$ subunit, which have a much lower affinity for nicotine, persist in the mesencephalic neurons of the mutant mice. These results indicate that $\beta 2$ -subunit-containing receptors contribute to most, if not all, of the response to nicotine of mesencephalic dopaminergic neurons.

As nicotine also affects dopamine metabolism and reuptake¹⁸, we investigated whether mutant animals showed any alterations in these systems. Extracellular levels of homovanillic acid (HVA) and dihydroxyphenyl acetic acid (DOPAC), the principal products of dopamine metabolism, were measured by high-performance liquid chromatography (HPLC) following *in vivo* microdialysis, and were found not to be significantly changed in the ventral striatum of $\beta 2$ -mutant mice compared with their wild-type siblings (results not shown). We tested the levels of the dopamine transporter, the principal means of dopamine reuptake, by using quantitative autoradiography with tritiated WIN35,428, a cocaine analogue¹⁹, and found no difference in WIN35,428 binding in the ventral and dorsal striatum (mutant: 36.2 ± 2.2 fmol g⁻¹ tissue; WT: 39.3 ± 2.8 fmol g⁻¹ tissue) or ventral mesencephalon (mutant: 45.2 ± 3.1 fmol g⁻¹ tissue; WT: 45.7 ± 7.4 fmol g⁻¹ tissue) of $\beta 2$ -mutant mice compared with their wild-type siblings. However, the activity of the transporter may be modified in the mutant mice, despite the lack of a change in equilibrium binding measurements.

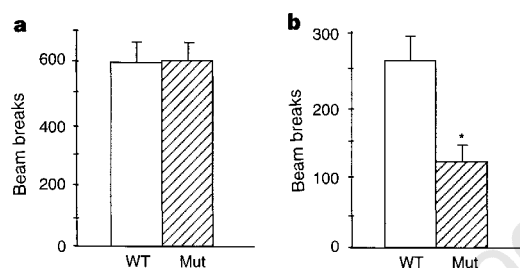


Figure 3 Locomotor activity of $\beta 2$ -mutant mice and their wild-type (+/+) siblings as measured in **a**, an unfamiliar environment, and **b**, in a familiar environment.

We also investigated whether mutant mice showed any difference in dopamine receptors and second messenger pathways involved in dopaminergic function. Quantitative receptor autoradiography using tritiated SCH23390 to detect D1 receptors²⁰ or tritiated raclopride to detect D2 receptors²¹, followed by computer-assisted image analysis, indicated that there was no significant difference in the levels of either D1 or D2 binding in ventral and dorsal striatum (D1: mutant, 410 ± 18 ; WT, 408 ± 14 . D2: mutant, 87.3 ± 9.4 ; WT, 88.1 ± 6.3 fmol g⁻¹ tissue) or SN and VTA (D1: mutant, 312 ± 15 ; WT, 315 ± 10 fmol g⁻¹ tissue) in mutant mice and their wild-type siblings. Dopamine-stimulated cyclase assays²² did not reveal any significant changes in basal levels of cyclic AMP (mutant: 0.14 ± 0.03 nM; WT: 0.13 ± 0.03 nM) and cyclase activity (mutant: 16.8 ± 2.2 nM cAMP; WT: 17.4 ± 1.7 nM cAMP), forskolin-stimulated adenylyl cyclase activity (mutant: 305.7 ± 33.4 nM cAMP; WT: 387.2 ± 46.0 nM cAMP), or dopamine-stimulated adenylyl cyclase activity (mutant: 35.2 ± 5.0 nM cAMP, maximum WT: 38.1 ± 2.7 nM cAMP, maximum), nor was tyrosine hydroxylase activity²³ altered (mutant: 30.7 ± 3.0 c.p.m. per μ g protein; WT: 34.3 ± 3.5 c.p.m. per μ g protein). The dose-response of dopamine-stimulated adenylyl cyclase activity is available as Supplementary Information. The lack of change in dopamine receptor levels and in the related second messenger systems in $\beta 2$ -mutant mice confirms that regulation of the mesostriatal dopamine system by high-affinity nAChRs occurs primarily on dopaminergic neurons rather than at their target cells in the striatum²⁴.

The change in the regulation of dopamine release by nicotine in the $\beta 2$ -mutant mice suggested that endogenous acetylcholine acting through nAChRs might also act physiologically in regulating dopamine transmission. The modulation of locomotor activity is a primary behavioural output of the mesostriatal dopamine system. In rodents, dopamine denervation reduces both exploratory activity and habituated locomotion⁸. To determine whether the high-affinity nAChR could be involved in the control of spontaneous locomotor activity, we tested the exploratory behaviour and the habituated locomotor activity of mice lacking the $\beta 2$ -subunit of their wild-type siblings in standard locomotor activity cages. Whereas the exploratory activity of mutant and wild-type animals was similar (Fig. 3a), the locomotor activity of the $\beta 2^{-/-}$ mice after habituation decreased by ~50% (Fig. 3b) compared to the wild-type siblings, indicating that endogenous acetylcholine might act through high-affinity nAChRs in the physiological regulation of locomotion.

The mesolimbic dopamine system is thought to be a key substrate for nicotine self-administration^{6,25}: rodents will press a lever to receive intravenous infusion of nicotine⁹, which results in activation of the terminal fields of the mesolimbic dopamine system²⁶. The role of the $\beta 2$ nicotinic receptor subunit in maintaining this operant response for nicotine was investigated in $\beta 2$ -mutant mice using an intravenous self-administration procedure^{27,28}. Mice were implanted with a jugular catheter and trained daily for cocaine

self-administration in operant chambers equipped with two nose-poke detectors, one active and one inactive. Once a stable baseline of cocaine responding in the active detector was reached, cocaine was replaced with nicotine at a very low dose, or with saline. As shown in Fig. 4, wild-type mice maintained significant specific operant response to nicotine during the 5 days following cocaine substitution. In wild-type mice, the pattern of nose-poke responses to nicotine, including some daily variations, resembled that in rats (ref. 28; and M. Epping-Jordan, personal communication). The performance of $\beta 2^{-/-}$ mice was significantly ($P < 0.05$) different from that of wild-type mice during the nicotine-substitution period, showing a decrease in nose-poke response (Fig. 4a; $P < 0.01$) and a clear-cut progressive reduction in discrimination between the active and inactive levers (Fig. 4b; $P < 0.01$). These effects resembled those of saline replacement in wild-type mice (Fig. 4), a substitution that does not maintain operant response. Thus, the ability of low doses of nicotine to maintain operant responding in mice trained to respond for cocaine in a substitution protocol was significantly attenuated in $\beta 2^{-/-}$ mice, suggesting that nicotine does not have a modulatory effect on the dopaminergic neurons of $\beta 2$ -mutant mice.

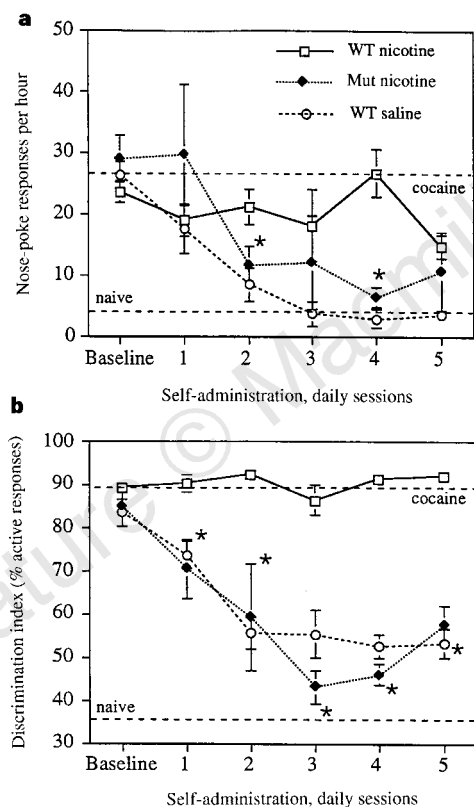


Figure 4 Intravenous self-administration of cocaine and nicotine in $\beta 2^{-/-}$ mice and their wild-type siblings with a history of cocaine self-administration. **a**, Number of active nose-poke responses per session. Dashed lines represent the average of the last 3 sessions of cocaine self-administration in all groups of mice ('cocaine'), or of spontaneous nose-poke behaviour in naive, sham-operated mice ('naive'). **b**, Per cent active responses: that is, the percentage of the number of active responses compared with the sum of active and inactive responses per session. Values of 50% indicate a lack of discrimination between active and inactive detectors. $\beta 2^{-/-}$ mice differed significantly from wild-type mice self-administering nicotine in either active nose-poke ($P < 0.05$) or discrimination index ($P < 0.01$), but did not differ from wild-type mice during saline-induced extinction (ANOVA). The increase in the variability of nose-poke response of $\beta 2^{-/-}$ mice on the first day of nicotine treatment was due to increased responding of some mice, usually interpreted as a transient overresponse to reinforcer devaluation. Asterisks represent *post-hoc* comparisons of the wild-type nicotine group and the $\beta 2^{-/-}$ group (* $P < 0.05$).

These results on nicotine-elicited dopamine release and the electrophysiological response of dopaminergic neurons to nicotine indicate that a $\beta 2$ -subunit-containing nAChR contributes to the effects of nicotine on the dopamine mesostriatal system. The changes in habituated locomotor activity of mutant mice in the absence of administered nicotine imply that the nicotinic regulation of the dopamine system may have consequences on the endogenous regulation of locomotion. Also, the changes in the self-administration of nicotine by $\beta 2$ -mutant mice indicate that a $\beta 2$ -subunit-containing nAChR is involved in mediating the reinforcing properties of nicotine, as determined by the attenuated response of $\beta 2$ -mutant mice when nicotine was substituted for cocaine in the self-administration protocol.

Experiments on nicotine-enhanced glutamatergic and cholinergic transmission in the developing chick brain using an $\alpha 7$ -type nAChR²⁹ indicate that low concentrations of nicotine can activate several different subtypes of nAChR, some of which do not contain a $\beta 2$ subunit. Many systems may converge to mediate the reinforcing properties of nicotine. Taking into account this heterogeneity, our results provide evidence that the lack of nicotine-mediated stimulation of the mesolimbic dopaminergic system is associated with attenuation of nicotine self-administration in $\beta 2$ -subunit-mutant mice. These data, combined with other evidence^{5,9} indicate that an increase in dopamine release mediated through $\beta 2$ -containing nAChRs is one of the primary substrates for the addictive properties of nicotine.

Methods

Microdialysis. Microdialysis and dopamine measurement were adapted from ref. 30. On the day of the experiment, mice were anaesthetized with a mixture of air and halothane (1.4% of halothane in the air flow of 1.51 min⁻¹) and placed in a stereotaxic frame. The skull was exposed and one hole was drilled to allow implantation of the microdialysis probe (CMA 11, o.d. 0.24 mm; 2-mm cuprofen dialysing membrane). The probe was implanted into the ventral striatum ($A = 0$, $L = 2.0$, $V = 4.5$) and was perfused at a flow rate of 2 μ l min⁻¹ with Ringer's solution. The site of probe implantation was chosen after extensive testing of rostral and caudal sites in ventral striatum (n , 5–6 animals per location), including core ($A = 0.7$, $L = 1.6$, $V = 4.8$) and shell ($A = 1.2$, $L = 1.0$, $V = 5.0$) of the nucleus accumbens. This location showed the most reliable and marked increase in nicotine-elicited dopamine release in our preparation. Collection of perfusate samples began 120 min after probe implantation, with 40- μ l samples being collected every 20 min. Nicotine was freshly dissolved in Ringer's solution and injected i.p. after collecting three stable samples that constituted the baseline. Upon completion of the experiment, mice were killed by an overdose of halothane and their brains were sectioned to verify the probe location. Dopamine, DOPAC and HVA concentrations were measured by HPLC with electrochemical detection. Limit of detection for dopamine was 2 fmol per sample.

Patch-clamp recording. Coronal slices through SN and VTA were obtained from 10–15-day-old mice for the nicotine dose-response curve and 1–2-month-old mice for experiments on discharge frequency. Patch-clamp recording was done as described³. To improve the quality of the slices, a step of intracardiac superfusion of the mice with an ice-cold sucrose solution was added before killing the animals at 1–2 months of age. The sucrose solution contained (in mM): sucrose 283, NaHCO₃ 26, KCl 1.3, CaCl₂ 2, MgCl₂ 2, D-glucose 10, and was bubbled with 95% O₂, 5% CO₂ (pH $\approx$ 7.2). For blocking experiments, dihydro- β -erythroidine (DH β E) was applied first in the bath and then together with nicotine.

Locomotor activity. A clear plastic cage 25.4 cm high $\times$, 45 cm long and 25.4 cm wide was used for all locomotor activity experiments. Movement was detected by eight infrared photobeam detectors and transducers set 1.5 cm above the floor of the apparatus and measured by a PC. Mice of 6–7 months old that were naive to the apparatus were used in novelty testing; activity was measured for 15 min, with time points being taken every 5 min ($n = 14$ per group). For locomotion in a familiar environment, activity was measured after repeated exposure to the apparatus for 60 min at intervals of 20 min ($n = 10$ per group). For all experiments, each mutant mouse was compared with a wild-

type (+/+) sibling of the same sex from the same litter. Locomotion experiments were run by an observer blind to the genotype of the animals being tested. Tests were run from 10 a.m. to 2 p.m. Data are presented as mean $\pm$ s.e.m. Statistical analysis was performed using ANOVA ($P < 0.05$ mutant compared with wild type).

Self-administration. Adult male mice (wild-type and mutant siblings of parents backcrossed 3 generations to C57 B1/6 inbred mice) were implanted with a silastic catheter in the jugular vein under halothane anaesthesia and tested in operant cages equipped with two nose-poke detectors, one active, the other inactive (V. Deroche *et al.*, manuscript in preparation). All mice were first trained with cocaine (0.8 mg kg⁻¹ per injection, delivered in 50 μ l per 2 s, with 20-s time-out period) under fixed-ratio (FR) 1 for 2–4 sessions, then under FR2 until stable baseline was reached. Spontaneous nose-poke behaviour, measured in naive, sham-operated mice when both detectors were inactive, was low and non-discriminatory (2.6 ± 0.6 per h), and did not differ between $\beta 2^{-/-}$ mice and wild-type mice (not shown). The baseline for each mouse was defined as 3 consecutive sessions with less than 30% deviation from the mean and at least 75% active-specific response. During 5 consecutive daily sessions, cocaine was replaced with nicotine (0.03 mg kg⁻¹ per injection, delivered in 50 μ l per 2 s, 20-s time-out period, under FR 2 schedule) in wild-type mice ($n = 5$) and $\beta 2^{-/-}$ mice ($n = 5$). In a second group of wild-type mice ($n = 5$), cocaine was replaced with saline, forcing the operant responding to extinction. Cocaine and nicotine bitartrate (Sigma) were freshly dissolved in saline before each experiment.

Equilibrium binding. Receptor autoradiography was done as described^{19–21}. Briefly, following 30 min preincubation in the appropriate buffer, 14- μ m coronal brain sections from wild-type and $\beta 2$ -mutant mice were incubated at room temperature for 120 min with 5 nM [³H]WIN35,428 (84.5 Ci mmol⁻¹; NEN)¹⁹ or for 60 min with either 1.5 nM SCH23390 (70 Ci mmol⁻¹; NEN)²⁰ or 3 nM raclopride (82.4 Ci mmol⁻¹; NEN)²¹. Slides were washed in ice-cold buffer twice for 1 min (WIN35,428) or for 5 min (SCH23390 or raclopride), and exposed to hyperfilm together with appropriate standards ([³H]microscale; Amersham). Nonspecific labelling was determined in the presence of 30 μ M cocaine for the dopamine transporter and 1 μ M (+)butaclamol for D1 and D2. Siblings of the same sex and litter were used for all experiments (n , 4–6 per group).

Cyclase and tyrosine hydroxylase assays. Dopamine-stimulated cyclase activity was measured in striatal homogenates from wild-type and $\beta 2$ -mutant siblings of the same sex and same litter by following published protocols²² and using the Amersham Biotrak scintillation proximity assay to determine cAMP levels. For each animal ($n = 6$ per group), the striatum from one side of the brain was used for dopamine-stimulated cyclase assays, and the striatum from the other side of the brain was used to measure tyrosine hydroxylase activity as described²³. Briefly, after homogenization, samples were incubated with [³H]tyrosine and tetrahydrobiopterin. End products were separated from unreacted [³H]tyrosine by treatment with activated charcoal. Results are reported as the number of counts incorporated minus background per μ g protein assayed.

Received 7 August; accepted 25 September 1997.

- Dani, J. A. & Heinemann, S. Molecular and cellular aspects of nicotine abuse. *Neuron* **16**, 905–908 (1996).
- Role, L. W. & Berg, D. K. Nicotinic receptors in the development and modulation of CNS synapses. *Neuron* **16**, 1077–1085 (1996).
- Picciotto, M. R. *et al.* Abnormal avoidance learning in mice lacking functional high-affinity nicotine receptor in the brain. *Nature* **374**, 65–67 (1995).
- Koob, G. F. Drugs of abuse: anatomy, pharmacology and function of reward pathways. *Trends Pharmacol. Sci.* **13**, 177–184 (1992).
- Stolerman, I. P., Mirza, N. R. & Shoaib, M. Nicotine psychopharmacology: Addiction, cognition and neuroadaptation. *Med. Res. Rev.* **15**, 47–72 (1995).
- Pontieri, F. E., Tanda, G., Orzi, F. & Di Chiara, G. Effects of nicotine on the nucleus accumbens and similarity to those of addictive drugs. *Nature* **382**, 255–257 (1996).
- Grady, S., Marks, M. J., Wonnacott, S. & Collins, A. C. Characterization of nicotinic receptor-mediated [³H]dopamine release from synaptosomes prepared from mouse striatum. *J. Neurochem.* **59**, 848–856 (1992).
- Clarke, P. B., Fu, D. S., Jakubovic, A. & Fibiger, H. C. Evidence that mesolimbic dopaminergic activation underlies the locomotor stimulant action of nicotine in rats. *J. Pharmacol. Exp. Ther.* **246**, 701–708 (1988).
- Corrigall, W. A., Franklin, K. B., Coen, K. M. & Clarke, P. B. The mesolimbic dopaminergic system is implicated in the reinforcing effects of nicotine. *Psychopharmacology (Berlin)* **107**, 285–289 (1992).
- Le Novère, N., Zoli, M. & Changeux, J.-P. Neuronal nicotinic receptor $\alpha 6$ subunit mRNA is selectively concentrated in catecholaminergic nuclei of the rat brain. *Eur. J. Neurosci.* **8**, 2428–2439 (1996).
- Reavill, C., Walther, B., Stolerman, I. P. & Testa, B. Behavioural and pharmacokinetic studies on nicotine, cytosine and lobeline. *Neuropharmacology* **29**, 619–624 (1990).
- Risinger, F. O. & Oakes, R. A. Nicotine-induced conditioned place preference and conditioned place aversion in mice. *Pharmacol. Biochem. Behav.* **51**, 457–461 (1995).

- Nisell, M., Nomikos, G. G. & Svensson, T. H. Systemic nicotine-induced dopamine release in the rat nucleus accumbens is regulated by nicotinic receptors in the ventral tegmental area. *Synapse* **16**, 36–44 (1994).
- Henningfield, J. E., Miyasato, K. & Jasinski, D. R. Cigarette smokers self-administer intravenous nicotine. *Pharmacol. Biochem. Behav.* **19**, 887–890 (1983).
- Reavill, C., Walther, B., Stolerman, I. P. & Testa, B. Behavioural and pharmacokinetic studies on nicotine, cytosine and lobeline. *Neuropharmacology* **29**, 619–624 (1990).
- Grace, A. A. & Onn, S. P. Morphology and electrophysiological properties of immunocytochemically identified rat dopamine neurons recorded *in vitro*. *J. Neurosci.* **9**, 3463–3481 (1989).
- Luetje, C. W. & Patrick, J. Both α - and β -subunits contribute to the agonist sensitivity of neuronal nicotinic acetylcholine receptors. *J. Neurosci.* **11**, 837–845 (1991).
- Izenwasser, S., Jacobs, H. M., Rosenberger, J. G. & Cox, B. M. Nicotine indirectly inhibits [³H]dopamine uptake at concentrations that do not directly promote [³H]dopamine release in rat striatum. *J. Neurochem.* **56**, 603–610 (1991).
- Kaufman, M. J., Spealman, R. D. & Madras, B. K. Distribution of cocaine recognition sites in monkey brain: I. *In vivo* autoradiography with [³H]CFT. *Synapse* **9**, 177–187 (1991).
- Dawson, T. M., Gehlert, D. R., McCabe, R. T., Barnett, A. & Wamsley, J. K. D1 dopamine receptors in the rat brain: a quantitative autoradiographic analysis. *J. Neurosci.* **6**, 2352–2365 (1986).
- Mansour, A. *et al.* Localization of dopamine D2 receptor mRNA and D1 and D2 receptor binding in the rat brain and pituitary: an *in situ* hybridization-receptor autoradiographic analysis. *J. Neurosci.* **10**, 2587–2600 (1990).
- Memo, M. *et al.* Dopaminergic inhibition of prolactin release and calcium influx induced by neurotensin in anterior pituitary is independent of cyclic AMP system. *J. Neurochem.* **47**, 1689–1695 (1986).
- Rheinhard, J. F., Smith, G. K. & Nichol, C. A. A rapid and sensitive assay for tyrosine-3-monooxygenase based upon the release of [³H]-H₂O and adsorption of [³H]-tyrosine by charcoal. *Life Sci.* **39**, 2185–2189 (1986).
- Blaha, C. D. *et al.* Modulation of dopamine efflux in the nucleus accumbens after cholinergic stimulation of the ventral tegmental area in intact, pedunculopontine tegmental nucleus-lesioned, and laterodorsal tegmental nucleus-lesioned rats. *J. Neurosci.* **16**, 714–722 (1996).
- Calabresi, P., Lacey, M. G. & North, R. A. Nicotinic excitation of rat ventral tegmental neurones *in vitro* studies by intracellular recording. *Br. J. Pharmacol.* **98**, 135–140 (1989).
- Merlo Pich, E. *et al.* Common neural substrates for the addictive properties of nicotine and cocaine. *Science* **275**, 83–86 (1997).
- Grahame, N. J., Phillips, T. J., Burkhart-Kasch, S. & Cunningham, C. L. Intravenous cocaine self-administration in the C57BL/6J mouse. *Pharmacol. Biochem. Behav.* **51**, 827–834 (1995).
- Tessari, M., Valerio, E., Chiamulera, C. & Beardsley, P. M. Nicotine reinforcement in rats with histories of cocaine self-administration. *Psychopharmacology* **121**, 282–283 (1995).
- McGehee, D. S., Heath, M. J., Gelber, S., Devay, P. & Role, L. W. Nicotine enhancement of fast excitatory synaptic transmission in CNS by presynaptic receptors. *Science* **269**, 1692–1696 (1995).
- Rimondini, R. *et al.* PE-60 increases dopamine but not GABA release in the dorsolateral neostriatum of the halothane anaesthetized rat. An *in vivo* microdialysis study. *Neurosci. Lett.* **177**, 53–57 (1994).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy form Mary Sheehan at the London editorial office of Nature.

Acknowledgements. We thank M. Memo and D. Uberti for help and advice with DA-stimulated cyclase assays; M. Epping-Jordan for help with self-administration experiments; P. Sattounet-Roché for technical assistance; L. Gold, F. Caine and C. Chiamulera for discussing the behavioural results; and E. Ratti, D. Trist and A. North (GlaxoWellcome) for supporting part of the project. This work was supported by the Collège de France, the Centre National de la Recherche Scientifique, the Association Française contre la Myopathie, the Council for Tobacco Research, Biomed and Biotech contract from the Commission of the European Communities, a grant from the Human Frontiers Science Program, a Roux grant from the Institut Pasteur for C.L., a young investigator award from NARSAD and a grant from NIDA to M.R.P.

Correspondence and requests for materials should be addressed to J.P.C.

Extreme Th1 bias of invariant V α 24J α Q T cells in type 1 diabetes

S. Brian Wilson^{*†}, Sally C. Kent^{†‡}, Kurt T. Patton[‡], Tihamer Orban[§], Richard A. Jackson[§], Mark Exley^{||}, Steven Porcell[¶], Desmond A. Schatz[#], Mark A. Atkinson[#], Steven P. Balk^{||}, Jack L. Strominger^{*} & David A. Hafler[‡]

^{*} Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138, USA

[‡] Center for Neurologic Diseases, Brigham and Women's Hospital; [§] Immunology Section, Joslin Diabetes Center; ^{||} Cancer Biology Program, Hematology/Oncology Division, Beth Israel Deaconess Medical Center; [¶] Lymphocyte Biology Section, Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital; Harvard Medical School, Boston, Massachusetts 02115, USA

[#] Departments of Pathology and Pediatrics, University of Florida, Gainesville, Florida 32610, USA

[†] These authors contributed equally to this work.

Type 1 diabetes (insulin-dependent diabetes mellitus, IDDM) is a disease controlled by the major histocompatibility complex (MHC) which results from T-cell-mediated destruction of pancreatic β -cells¹. The incomplete concordance in identical twins and the presence of autoreactive T cells and autoantibodies in individuals who do not develop diabetes suggest that other

abnormalities must occur in the immune system for disease to result^{2,3}. We therefore investigated a series of at-risk non-progressors and type 1 diabetic patients (including five identical twin/triplet sets discordant for disease). The diabetic siblings had lower frequencies of CD4⁺CD8[−] Vα24JαQ⁺ T cells compared with their non-diabetic sibling. All 56 Vα24JαQ⁺ clones isolated from the diabetic twins/triplets secreted only interferon (IFN)-γ upon stimulation; in contrast, 76 of 79 clones from the at-risk non-progressors and normals secreted both interleukin (IL)-4 and IFN-γ. Half of the at-risk non-progressors had high serum levels of IL-4 and IFN-γ. These results support a model for IDDM in which Th1-cell-mediated tissue damage is initially regulated by Vα24JαQ⁺ T cells producing both cytokines; the loss of their capacity to secrete IL-4 is correlated with IDDM.

The discovery of Th1 and Th2 subsets of CD4⁺ T cells has helped to explain the cellular basis for the diversity of T-cell responses in autoimmunity⁴. Th1 cells promote inflammatory cellular immune responses and are biased towards secretion of IFN-γ, tumour-necrosis factor (TNF)-β and IL-2. Th2 cells are biased towards secretion of interleukins 4, 5, 6, 10 and 13, induce humoral immunity, and inhibit Th1 responses. Lymphocyte cytokine production in type 1 diabetes is known to exhibit a bias towards the Th1 cytokine IFN-γ (ref. 5), but the cellular mechanisms integrating the drive to Th1 or Th2 effector cell differentiation are poorly understood. In the mouse, one mechanism by which Th2 rather than Th1 T-cell bias may be promoted is by activation of invariant (no N/P nucleotide additions within the CDR3 of the *TCRA* gene) Vα14Jα281 TCR⁺ NK 1.1⁺ T cells capable of early secretory bursts of IL-4 and IFN-γ. The ligand for this family of T cells is CD1.1 on the surface of antigen-presenting cells^{6,7}.

CD4⁺CD8[−] T cells in humans expressing the invariant Vα24JαQ T-cell antigen receptor (TCR) which has close sequence homology to the murine Vα14Jα281 TCR have been described⁸. To determine whether there could be a relationship between the number of circulating CD4⁺CD8[−] Vα24JαQ⁺ T cells and type 1 diabetes, we did a frequency analysis on a set of type 1 diabetic discordant monozygotic twins and triplets. The numbers of circulating CD4⁺CD8[−] Vα24JαQ⁺ T cells in diabetes-free twins/triplets were compared with those present in their siblings with disease. The percentage of circulating invariant CD4⁺CD8[−] Vα24JαQ T cells could be determined by multiplying the frequency of invariant Vα24JαQ sequences present in the total CD4⁺CD8[−] Vα24⁺ population by the percentage of CD4⁺CD8[−] Vα24⁺ T cells, as measured by flow cytometric analysis (Table 1). No CD4⁺CD8[−] Vα24JαQ⁺ T cells were detected in three diabetics, despite at least three sorting attempts for each subject. The percentage of CD4⁺CD8[−] Vα24JαQ⁺ T cells in a previously disease-free diabetic twin (patient 6A; Table 1), studied during the week of IDDM diagnosis, was similar to that in the long-term IDDM twin and in the other diabetics. In all sets of family pairings, the IDDM sibling had markedly lower percentages of CD4⁺CD8[−] Vα24JαQ⁺ T cells ($P = 0.015$, paired sign test using only the discordant twins/triplets data).

To determine whether human Vα24JαQ⁺ T cells were functionally altered in type 1 diabetics and those at risk for the disease, single CD4⁺/CD8[−] mononuclear cells expressing Vα24⁺ TCR were cloned. The initial analysis was carried out on clones generated from the IDDM non-progressing member of a sibling pair, subject 7A (Table 1). All clones expressed the invariant Vα24JαQ junctional sequences conserving the germ-line-encoded amino acids Vα24 (-CVVS-) and JαQ (:DRGST-). Eight of ten clones were Vβ11⁺ and two were Vβ13⁺. All of the clones were CD4⁺ and uniformly negative when stained for CD8 β-chain. Surface expression of CD8α⁺ appeared to reflect the activation state, as staining for this marker reverted to negative 2–3 weeks post-stimulation. All T-cell clones expressed the human homologue of the murine NK1.1 molecule, NKR-P1A (ref. 9), and the C-type lectins encoded by the natural killer (NK) locus, CD69 and CD94 (data not shown).

CD1d restriction was assessed by co-cultivating Vα24JαQ⁺ T-cell clones with C1R cells transfected with a CD1d or control expression vector¹⁰. A T-cell clone (4.2) with a non-invariant TCR α-chain (Vα24N3Jα6) was included as a negative control. All T-cell clones, except 3.5, 3.8 and the control clone 4.2, specifically proliferated in response to the CD1d transfectant of C1R (Fig. 1a). All of the clones except 3.5 and 4.2 secreted IL-4 and IFN-γ in a CD1d-specific manner (Fig. 1b, c). Clone 3.5 secreted only IFN-γ in response to CD1d (Fig. 1). The fine specificity of the clones for CD1d was tested by using C1R targets transfected with CD1a, CD1c, CD1d or vector alone. Only CD1d-expressing target cells specifically stimulated each of the CD4⁺CD8[−] Vα24JαQ⁺ clones, as assessed by IL-4 and IFN-γ secretion (Fig. 2).

A panel of Vα24JαQ⁺ T-cell clones was then raised from: (1) the twins/triplets discordant for type 1 diabetes (Table 1); (2) from an additional four at-risk non-progressors with raised serum IL-4 levels (see below); and (3) two haplotype (DR3/DR2 and DR4/DRX)-matched normal controls. Twenty-five out of 28 clones raised from the at-risk non-progressors among the discordant twins/triplets secreted both IL-4 and IFN-γ (>10 pg ml^{−1}) on stimulation with anti-CD3 (Fig. 3a). The other three clones produced only IFN-γ. Unlike the other non-progressing twins, only one clone from triplet 1A secreted moderate amounts of IL-4 when stimulated. Only a single attempt to generate clones from this subject was made owing to subsequent entry into a clinical trial. All of the 56 clones raised from the diabetic twins/triplets secreted only IFN-γ with anti-CD3 stimulation, and diabetic twins 4B and 5B had no identifiable CD4⁺CD8[−] Vα24JαQ⁺ T cells (Table 1). There was no difference in the proliferative response to anti-CD3 between the clones raised from diabetics or other subjects (data not shown). The new onset type 1 twin 6A (Table 1) had 9/9 CD4⁺CD8[−] Vα24JαQ⁺ T-cell clones that secreted only IFN-γ (Fig. 3b). This suggested that the Th1 phenotype seen in the new-onset twin was not related to duration of diabetes but occurred before, or concurrently with, the

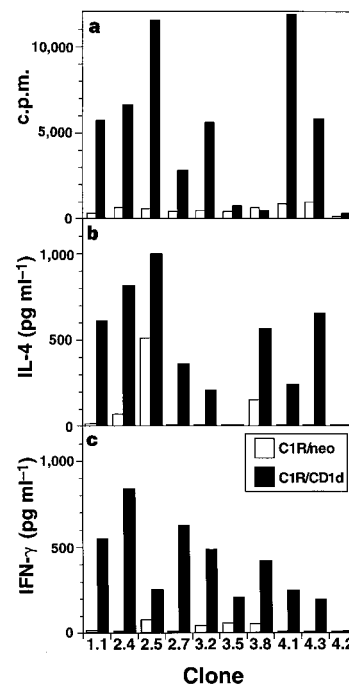


Figure 1 CD4⁺CD8[−] Vα24JαQ T-cell clones respond to C1R/CD1d transfectants specifically. Vα24JαQ T cell clones (1.1, 2.4, 2.5, 2.7, 3.2, 3.5, 3.8, 4.1 and 4.3) and control clone 4.2 (Vα24N3Jα6), all at 5×10^4 per well, were stimulated with fixed C1R/CD1d or C1R/neo transfectant cells at 5×10^4 per well. PMA (1 ng ml^{−1}) was added. **a**, Proliferation was measured by tritiated thymidine incorporation at 72 h; **b**, secreted IL-4, and **c**, IFN-γ were assayed by ELISA at 48 h. All clones were from subject 7A. One of three representative experiments is shown.

onset of overt disease. As individuals discordant for the development of type 1 diabetes were genetically identical, the question remains as to what environmental factor(s), if any, may have triggered the decreased frequency and cytokine shift of the V α 24J α Q T cells.

An additional set of 33 clones was generated from four at-risk non-progressors (see below and Fig. 4), and 18 clones were raised from MHC haplotype-matched controls (Fig. 3b). Clones raised from these subjects were phenotypically similar to the diabetes-free twins and a series of invariant V α 24J α Q⁺ T-cell clones previously described^{10,11}. Thus, all V α 24J α Q T-cell clones raised from type 1 patients showed an extreme Th1 bias, making them incapable of providing the IL-4 necessary for initiation of Th2 responses. In fact,

unopposed IFN- γ secretion could augment or initiate a Th1-dominated cellular attack on pancreatic β -cells^{12,13}.

Our functional studies on V α 24J α Q⁺ T cells from discordant twins/triplets, at-risk non-progressors and controls suggested that these two groups had polarized cell-mediated immune responses. We therefore tested serum IL-4 and IFN- γ in 14 at-risk IDDM non-progressors who had remained IDDM-free despite a 50% risk of developing diabetes during the study¹⁴. This cohort was defined by having remained healthy despite five or more years follow-up after diagnosis of type 1 diabetes in a first-degree relative and being positive for islet-cell antibodies (ICA⁺) with any two of the following autoantibodies: anti-GAD, anti-IA2 or anti-insulin autoantibodies. Seven of 14 type 1 non-progressors had markedly raised levels of serum IL-4 (Fig. 4), six of whom also had raised IFN- γ (0.2–35 ng ml⁻¹). Despite the increase in cytokines in the serum from

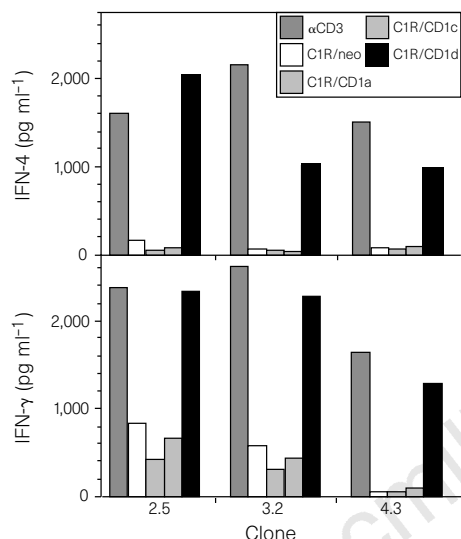


Figure 2 Specificity of three CD4⁺CD8⁻ V α 24J α Q T-cell clones for CD1 isoforms. Clones 2.5, 3.2 and 4.3 were co-cultivated with the following C1R transfectants: C1R/neo, C1R/CD1a, C1R/CD1c, and C1R/CD1d, as for Fig. 1: top, secreted IL-4; bottom, secreted IFN- γ . In addition, clones were activated with plate-bound anti-CD3 or immunoglobulin control.

Table 1 Frequency of CD4⁺CD8⁻ V α 24J α Q T cells from IDDM and disease-free siblings

Twins/triplets	DN (%) [*]	% DN V α 24 ⁺ in total lymphocytes	V α 24J α Q DN sequence frequency	V α 24J α Q (%)
1A/IL-4 ⁺	0.74	0.04	20/22	0.036
1B/IDDM	0.95	0.01	10/19	0.005
1C/IDDM	0.76	0.04	9/22	0.016
2A	2.1	0.37	9/10	0.33
2B/IDDM	3.1	0.025	31/31	0.025
3A	1.1	0.04	8/12	0.027
3B/IDDM	1.89	0.01	5/15	0.003
4A	1.21	0.02	4/13	0.006
4B/IDDM	0.31	0.006	0†	0
5A	0.58	0.06	8/12	0.04
5B/IDDM	0.98	0	0†	0
6A/new IDDM	0.89	0.03	7/26	0.008
6B/IDDM	2.62	0.03	8/23	0.01
Brother/sister				
7A/IL-4 ⁺	2.54	0.03	8/12	0.017
7B/IDDM	1.08	0.005	0/18	0

The frequency of V α 24J α Q TCR sequences was determined by sorting all CD4⁺CD8⁻ α β TCR⁺ T cells, amplifying all V α 24 transcripts and sequencing the TCR CDR3 region; the percentage of cells that were invariant CD4⁺CD8⁻ V α 24J α Q in total mononuclear cells was calculated by multiplying the sequence frequency by the CD4⁺CD8⁻ V α 24⁺ percentage of total mononuclear cells, as determined by flow cytometry. IL-4⁺ indicates a subject with high serum IL-4. ^{*} DN, CD4⁺CD8⁻. [†] No V α 24 PCR products were detected in three attempts.

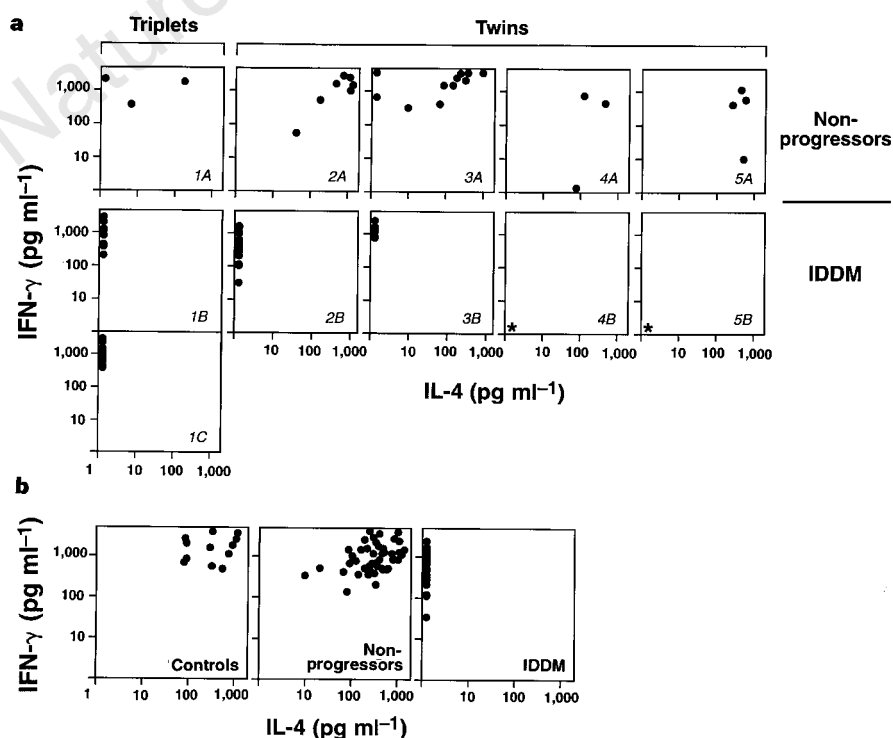


Figure 3 IL-4 and IFN- γ secretion profiles of CD4⁺CD8⁻ V α 24J α Q T-cell clones raised from monozygotic twins and triplets discordant for IDDM. **a**, Plate-bound anti-CD3 or control immunoglobulin was used to stimulate individual clones and secreted IL-4 and IFN- γ was assayed at 4 h. The pattern of cytokine secretion was similar at 24 h. The CDR3 T-cell receptor sequences were determined and the surface phenotype was confirmed by flow cytometry for every clone (data not shown). Asterisks, twins with IDDM (4B and 5B) had no detectable V α 24J α Q T cells in two attempts to sort from 20 million PBMCs (Table 1). **b**, CD4⁺CD8⁻ V α 24J α Q T-cell clones from 4 IDDM non-progressing subjects (Fig. 4), two control (DR2/DR3 and DR4/DRX) and five IDDM patients (triplets 1B and 1C, twins 2B and 3B, and new-onset IDDM twin 6A), were assayed as described in **a**.

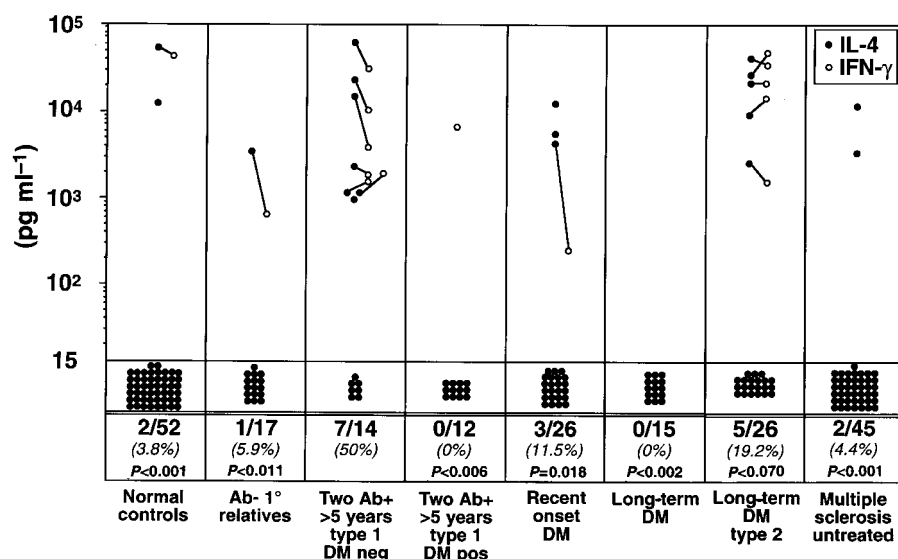


Figure 4 Serum IL-4 and IFN- γ in type 1 diabetes. Sera were collected and assayed for IL-4 and IFN- γ by capture ELISA, the detection minimum for which was 15 pg ml⁻¹. Raised cytokines were detected in 7/14 members of the at-risk IDDM non-progressor cohort. Lines connecting symbols indicate samples in which both IL-4 and IFN- γ were detected; no IFN- γ was detected in six subjects with raised IL-4. **P*, Two-tailed Fisher exact test comparing high IL-4 frequency in each cohort with at-risk IDDM non-progressors.

seven of fourteen non-progressors, all seven IL-4⁺ individuals have remained healthy with no evidence of chronic infectious or atopic/allergic illnesses. Both cytokines in the remaining seven were below the detection limit of our enzyme-linked immunosorbent assay (ELISA; 0.015 ng ml⁻¹). Five of 14 individuals in this group were found to have the strongly protective MHC allele DQB1*0602 and therefore are not at the same risk of progression as the remaining nine members of this cohort¹⁵. Three of these five individuals had raised serum IL-4 and IFN- γ levels, in contrast to our finding that IL-4 was undetectable in the serum (before or after diagnosis of type 1 diabetes) in 12 individuals with identical autoantibody status who developed IDDM after five or more years of follow up (Fig. 4).

Raised cytokines were also detected in archival serum samples obtained from 3/23 individuals at the time of diagnosis of type 1 diabetes and in 5/26 type-2 diabetics who did not have autoantibodies or a family history of type 1 diabetes (Fig. 4). When compared with normals, antibody-positive first-degree relatives, recent-onset diabetics, long-term diabetics (IDDM >2 years), autoantibody-negative first-degree relatives or untreated patients with multiple sclerosis (MS), the frequency of serum IL-4⁺ individuals was significantly raised in the non-progressor cohort (Fig. 4). The authenticity of the IL-4 detected was confirmed independently by using another set of ELISA antibodies, by binding to soluble recombinant IL-4 receptor produced in insect cells, and by western blot (data not shown).

Our results demonstrate a relationship between elevated serum IL-4 levels and resistance to the progression of an autoimmune disorder. Prolonged hyperglycaemia as an explanation for the absence of IL-4 in type 1 diabetics is less likely because IL-4 was detected in the serum of type 2 diabetics. Increased IL-4 was not an absolute predictor of IDDM resistance as only half of the resistant cohort had raised serum IL-4, as did 3/23 diabetics at or near the time of diagnosis.

In the non-obese diabetic (NOD) mouse, there is evidence that IL-4 exerts a dominant-negative effect on the progression to IDDM¹⁶⁻¹⁸. Differentiation of T cells into IL-4 secreting Th2 effector cells requires IL-4 priming⁴. Although this proposed function for NK1.1⁺ T cells was not obligatory for all Th2 immune responses^{6,19,20}, T-cell IL-4 secretion was markedly diminished in a CD1 knockout background²⁰⁻²². Fewer NK 1.1⁺ T cells were found to be present and were less frequent before the onset of disease in several murine methods of autoimmunity^{6,7,23-25}. In these models, autoimmunity was accelerated by depletion of NK 1.1⁺ T cells and delayed by generating mice transgenic for the V α 24J α 281 TCR. Diabetes was also prevented in the NOD mouse by adoptive transfer of a population harbouring the NK1.1-like class of T cell²⁶.

In summary, type 1 diabetes is associated with an extreme Th1 phenotype for V α 24J α Q⁺ T cells and a decrease in their circulating frequency. Our results suggest that there is a strong link between V α 24J α Q⁺ T cells and human type 1 diabetes; this indicates that they may be functionally related to the resistance or progression of this autoimmune disease in humans. □

Methods

Antibodies and phenotypic analysis of T cells. Flow cytometry experiments were done on FACScaliber and FACS Vantage instruments (Becton Dickinson). Monoclonal antibody (mAb) DX1 was a gift from L. Lanier; anti-CD4, anti-CD8 and anti-panTCR were from Becton Dickinson; anti-V α 24, anti-CD8 β , anti-CD56, anti-CD16 and anti-p58CD158 (NK workshop mAbs GL183 and EB6) were from Immunotech; anti-CD69 and anti-CD94 were from Pharmingen.

CDR3 TCR sequencing. Total CD4⁺CD8⁻ V α 24J α Q CDR3 sequences were amplified by reverse transcription followed by polymerase chain reaction (RT-PCR) using V α 24 and constant-region-specific primers as described⁸, and cloned using a Stratagene pCR-Script kit. Individual T-cell clones TCR transcripts were amplified by RT-PCR. Sequences of the plasmid and PCR DNA products were determined directly on an ABI 373A Automated DNA Sequencer.

Cell culture and cytokine assay. Single-cell sorts were grown with a mixture of irradiated (5,000 rad) allogeneic feeders at 50,000 per well and 721.221 cells at 5,000 per well and supplemented with 1 μ g ml⁻¹ PHA-P, IL-2 and IL-7 each at 10 U ml⁻¹, then propagated as described²⁷. Clones positive for V α 24 and NKR-P1A by flow cytometry and a V α 24J α Q CDR3 TCR sequence were assayed for cytokine secretion. Cells were stimulated (25,000 per well) with plate-bound anti-CD3 (1 μ g ml⁻¹; Immunotech) or control isotype antibody (Sigma) for 4, 8 or 24 h. Supernatants were assayed for IL-4 and IFN- γ by capture ELISA, and after 24 h, 1 μ Ci per well of [³H]thymidine was added and incorporation measured as described²⁷.

CD1 restriction. Restriction experiments using CD1 isoform (CD1a, CD1c, CD1d and pSR α -neo vector alone) transfected C1R cells were done as described¹⁰.

Received 11 June; accepted 17 September 1997.

1. Davies, J. L. *et al.* A genome-wide search for human type-1 diabetes susceptibility genes. *Nature* **371**, 130–136 (1994).
2. Verge, C. F. *et al.* Late progression to diabetes and evidence for chronic β -cell autoimmunity in identical twins of patients with type 1 diabetes. *Diabetes* **44**, 1176–1179 (1995).
3. Roep, B. O. T cell responses to autoantigens in IDDM. The search for the holy grail. *Diabetes* **45**, 1147–1156 (1996).
4. Abbas, A. K., Murphy, K. M. & Sher, A., Functional diversity of helper T lymphocytes. *Nature* **383**, 787–793 (1996).
5. Kallmann, B. A. *et al.* Systematic bias of cytokine production toward cell-mediated immune regulation in IDDM and toward humoral immunity in Graves' disease. *Diabetes* **46**, 237–243 (1997).
6. Bendelac, A., Rivera, M. N., Park, H.-S. & Roark, J. H. Mouse CD1-specific NK1 T cells: development, specificity, and function. *Annu. Rev. Immunol.* **15**, 535–562 (1997).
7. Vicari, A. P. & Zlotnik, A. Mouse NK1.1⁺ T cells: a new family of T cells. *Immunol. Today* **17**, 71–76 (1996).

8. Porcelli, S., Yockey, C. E., Brenner, M. B. & Balk, S. P. Analysis of T cell antigen receptor (TCR) expression by human peripheral blood CD4⁺8⁺α/β T cells demonstrates preferential use of several Vβ genes and an invariant TCR α chain. *J. Exp. Med.* **178**, 1–16 (1993).
9. Lanier, L. L., Chang, C. & Phillips, J. H. Human NKR-PIA. A disulfide-linked homodimer of the C-type lectin superfamily expressed by a subset of NK and T lymphocytes. *J. Immunol.* **153**, 2417–2428 (1994).
10. Exley, M., Garcia, J., Balk, S. P. & Porcelli, S. Requirements for CD1d recognition by human invariant Va24⁺ CD4⁺CD8⁺ T cells. *J. Exp. Med.* **186**, 1–11 (1997).
11. Davodeau, F. *et al.* Close phenotypic and functional similarities between human and murine αβ cells expressing invariant TCR α-chains. *J. Immunol.* **158**, 5603–5611 (1997).
12. von Herrath, M. G. & Oldstone, M. B. A. Interferon-γ is essential for destruction of β cells and development of insulin-dependent diabetes mellitus. *J. Exp. Med.* **185**, 531–539 (1997).
13. Denkers, E. Y., Scharton-Kersten, T., Barbieri, S., Caspar, P. & Sher, A. A role for CD4⁺NK1.1⁺ T lymphocytes as major histocompatibility complex class II independent helper cells in the generation of CD8⁺ effector function against intracellular infection. *J. Exp. Med.* **184**, 131–139 (1996).
14. Verge, C. F. *et al.* Prediction of type 1 diabetes in first-degree relatives using a combination of insulin, GAD, and ICA512bdc/IA-2 autoantibodies. *Diabetes* **45**, 926–933 (1996).
15. Pugliese, A. *et al.* HLA-DQB1*0602 is associated with dominant protection from diabetes even among islet cell antibody-positive first-degree relatives of patients with IDDM. *Diabetes* **44**, 608–613 (1995).
16. Rapoport, M. J. *et al.* Interleukin 4 reverses T cell proliferative unresponsiveness and prevents the onset of diabetes in nonobese diabetic mice. *J. Exp. Med.* **178**, 87–99 (1993).
17. Mueller, R., Krahl, T. & Sarvetnick, N. Pancreatic expression of interleukin-4 abrogates insulinitis and autoimmune diabetes in nonobese diabetic (NOD) mice. *J. Exp. Med.* **184**, 1093–1099 (1996).
18. Fox, C. J. & Danska, J. S. IL-4 expression at the onset of islet inflammation predicts nondestructive insulinitis in nonobese diabetic mice. *J. Immunol.* **158**, 2414–2424 (1997).
19. Brown, D. R. *et al.* B2-microglobulin-dependent NK1.1⁺ T cells are not essential for T helper cell 2 immune responses. *J. Exp. Med.* **184**, 1295–1304 (1996).
20. Smiley, S. T., Kaplan, M. H. & Grusby, M. J. Immunoglobulin E production in the absence of interleukin-4-secreting CD1-dependent cells. *Science* **275**, 977–979 (1997).
21. Chen, Y.-H., Chiu, N. M., Mandal, M., Wang, N. & Wang, C.-R. Impaired NK1⁺ T cell development and early IL-4 production in CD1-deficient mice. *Immunity* **6**, 459–467 (1997).
22. Mendiratta, S. K. *et al.* CD1d1 mutant mice are deficient in natural T cells that promptly produce IL-4. *Immunity* **6**, 469–477 (1997).
23. Takeda, K. & Dennert, G. The development of autoimmunity in C57BL/6 *lpr* mice correlates with the disappearance of natural killer type 1-positive cells: evidence for their suppressive action on bone marrow stem cell proliferation, B cell immunoglobulin secretion, and autoimmune symptoms. *J. Exp. Med.* **177**, 155–164 (1993).
24. Mieza, M. A. *et al.* Selective reduction of Vα14⁺ NK T cells associated with disease development in autoimmune-prone mice. *J. Immunol.* **156**, 4035–4040 (1996).
25. Gombert, J.-M. *et al.* IL-7 reverses NK1⁺ T cell-defective IL-4 production in the non-obese diabetic mouse. *Int. Immunol.* **8**, 1751–1758 (1996).
26. Baxter, A. G., Kinder, S. J., Hammond, K. J. L., Scollay, R. & Godfrey, D. I. Association between αβTCR⁺CD4⁺CD8⁺ T cell deficiency and IDDM in NOD/Lt mice. *Diabetes* **46**, 572–582 (1997).
27. Fukaura, H. *et al.* Induction of circulating myelin basic protein and proteolipid-specific transforming growth factor-β-secreting Th3 cells by oral administration of myelin in multiple sclerosis patients. *J. Clin. Invest.* **98**, 70–77 (1996).

Acknowledgements. We thank the DFCE flow cytometry facility for cell sorting, T. Smith for patient support, J. Orov for statistical analysis and A. LaMotte for clerical support. We wish to specifically thank the patients for their generosity without which these studies could not have been performed. These studies were supported by grants from the NIAID of the NIH.

Correspondence and requests for materials should be addressed to D.A.H.

Drosophila Shaking-B protein forms gap junctions in paired *Xenopus* oocytes

Pauline Phelan, Lucy A. Stebbings, Richard A. Baines[†], Jonathan P. Bacon, Jane A. Davies & Chris Ford^{*}

Sussex Centre for Neuroscience and ^{*}Department of Genetics and Development, School of Biological Sciences, University of Sussex, Brighton BN1 9QG, UK

In most multicellular organisms direct cell–cell communication is mediated by the intercellular channels of gap junctions. These channels allow the exchange of ions and molecules that are believed to be essential for cell signalling during development and in some differentiated tissues. Proteins called connexins, which are products of a multigene family, are the structural components of vertebrate gap junctions^{1,2}. Surprisingly, molecular homologues of the connexins have not been described in any invertebrate. A separate gene family, which includes the *Drosophila* genes *shaking-B* and *l(1)ogre*, and the *Caenorhabditis elegans* genes *unc-7* and *eat-5*, encodes transmembrane proteins with a predicted structure similar to that of the connexins^{3–9}. *shaking-B* and *eat-5* are required for the formation of functional gap junctions^{8,10}. To test directly whether Shaking-B is a channel protein, we expressed

it in paired *Xenopus* oocytes. Here we show that Shaking-B localizes to the membrane, and that its presence induces the formation of functional intercellular channels. To our knowledge, this is the first structural component of an invertebrate gap junction to be characterized.

The *shaking-B* (*shak-B*) locus was identified in *Drosophila* in two independent screens for behavioural mutants^{11,12}. Two transcripts, referred to as *shak-B(neural)* (formerly *Passover*) and *shak-B(lethal)*, which share five 3' exons, encode proteins of relative molecular mass 43,000 to 44,000 (*M_r* 43K–44K)^{3–5}. Both are predicted to have four transmembrane domains, two extracellular loops, and cytoplasmic amino and carboxy termini³. By comparing dye-coupling in wild-type and mutant flies, we have shown that the product of *shak-B(neural)* is essential for the function of the gap junctions at electrical synapses in the giant fibre system, the pathway that subserves the insect's escape response¹⁰, and in some embryonic somatic muscle (R.A.B. *et al.*, unpublished data). These studies in *Drosophila* do not tell us whether Shak-B proteins are integral channel components of gap junctions or accessory proteins necessary for gap junctions to function. To distinguish between these possibilities we expressed *shak-B* RNAs in the *Xenopus* oocyte system, which has been used extensively to characterize the channel-forming ability of the vertebrate connexins¹³.

RNAs encoding Shak-B(neural) and Shak-B(lethal) were micro-injected, either individually or together, into the vegetal hemisphere of single oocytes. We first checked whether the oocytes efficiently

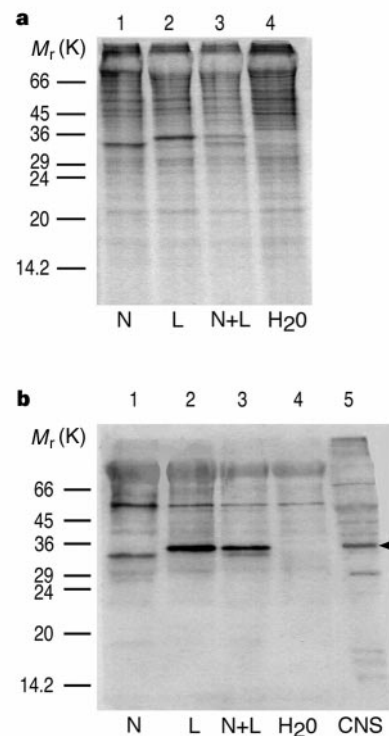


Figure 1 *shak-B* RNA is translated in the *Xenopus* oocyte expression system. **a**, Total radiolabelled proteins in membranes (single cell equivalent) prepared from oocytes 24 h after injection of [³⁵S] methionine (0.3 μCi) and 10 ng of *shak-B(neural)* RNA (N, lane 1), *shak-B(lethal)* RNA (L, lane 2), both RNAs (N + L, lane 3), or water (H₂O, lane 4). Bands unique to the RNA-injected cells (lanes 1–3) are Shak-B proteins. **b**, Western blot of proteins in membranes (half-cell equivalent) of RNA or water-injected oocytes (lanes 1–4, as in **a**) and in a homogenate of adult *Drosophila* nervous systems (CNS, lane 5). The blot was probed with Shak-B antiserum¹⁰. Labelled proteins (lanes 1–3) correspond in size to the proteins uniquely detected in **a** (lanes 1–3), and to native *Drosophila* Shak-B protein (lane 5, arrowhead). Positions of *M_r* markers are indicated on the left.

[†] Present address: Zoology Department, University of Cambridge, Cambridge CB2 3EJ, UK.

translate Shak-B proteins and insert them into membranes. [^{35}S]-methionine was microinjected after the RNA to label newly synthesized proteins. In denaturing polyacrylamide gels of oocyte membrane fractions (Fig. 1a), we detected strong bands at 34K and 36K in oocytes injected with *shak-B(neural)* (Fig. 1a, lane 1) and *shak-B(lethal)* (Fig. 1a, lane 2), respectively. Two bands at the appropriate molecular weights are visible in oocytes co-injected with the *shak-B* RNAs; levels of the individual proteins were consistently lower because the RNAs presumably compete for the translational machinery (Fig. 1a, lane 3). Labelled bands of the same molecular weights were not present in water-injected controls (Fig. 1a, lane 4), indicating that these are the products of exogenous *shak-B* RNA. Densitometric scanning of the relevant bands (Fig. 1a) indicated that the levels of Shak-B(neural) and Shak-B(lethal) were almost identical; average optical densities were 0.85 and 0.88, respectively (Fig. 1a, lane 3). Note that each form of the protein contains the same number of methionine residues.

To further confirm that the labelled bands uniquely detected in the RNA-injected oocytes (Fig. 1a) are Shak-B proteins, we probed

western blots with Shak-B peptide antiserum (Fig. 1b). The antibody recognizes a 34K protein in the membrane fraction from cells injected with *shak-B(neural)* RNA, a 36K protein in *shak-B(lethal)*-injected cells, and both in oocytes that received both RNAs. Shak-B proteins migrate anomalously in polyacrylamide gels. Apparent M_r values of 34K–36K for the expressed proteins (Fig. 1), although lower than their true M_r deduced from the amino-acid sequences^{3–5}, correspond to the apparent size of native Shak-B present in the nervous system of adult *Drosophila* (Fig. 1b, lane 5). We consistently found stronger antibody labelling of the product of *shak-B(lethal)* than *shak-B(neural)* RNA (Fig. 1b). This was unexpected because the C-terminal peptide to which the antiserum was raised is common to both proteins¹⁰. There is no simple explanation for this anomaly. As a precaution, we resequenced the 3' end of the *shak-B(neural)* construct and verified that it is complete.

To determine whether *shak-B* RNAs induced the formation of intercellular channels, RNA-injected oocytes were manipulated into contact with the vegetal membranes apposed. After 24 h we tested for the presence of electrical coupling using a double voltage-clamp procedure. Water-injected cells were included in all experiments to control for artefactual coupling or to measure coupling resulting

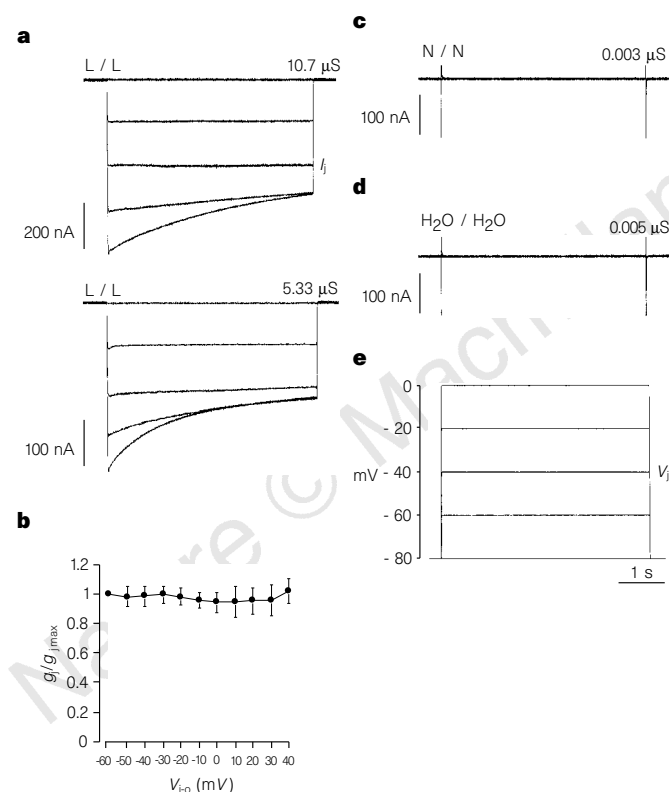


Figure 2 Shak-B(lethal) but not Shak-B(neural) protein forms voltage-sensitive intercellular channels in paired oocytes. **a, c–e**, Both oocytes of a pair were clamped to -80 mV . From this holding potential (V_h), sequential depolarizing voltage steps (V_j) were applied to one cell (cell 1; **e**) to generate a transjunctional voltage, V_j ($V_j = V_1 - V_h$). The current, I_j ($I_j = -I_2$) required to maintain the neighbouring cell (cell 2; **a, c, d**) at the holding potential was recorded. Junctional conductance (g_j) for each pair is I_j/V_j . **b**, Both oocytes of a pair were simultaneously clamped to equal potentials to establish inside-outside voltage (V_{i-o}) over the range -60 to $+40\text{ mV}$. At each potential a 10 mV hyperpolarizing pulse was delivered to cell 1 to measure g_j . Values ($\pm$ s.d. for $n = 13$ (-60 to 20 mV), $n = 7$ (30 mV), $n = 4$ (40 mV)) were normalized to g_j at -60 mV ($g_{j\text{ max}}$). **a**, Cell pairs injected with *shak-B(lethal)* RNA (0.5 ng , L/L) form functional homotypic channels; voltage steps delivered to cell 1 produce large I_j in cell 2. The g_j depends on V_j ; at $V_j \geq 60\text{ mV}$ the channels partly close so that steady-state I_j is reduced. **b**, g_j in L/L pairs is independent of V_{i-o} . **c, d**, In cell pairs injected with *shak-B(neural)* RNA (0.5 ng , N/N; **c**) or water ($\text{H}_2\text{O}/\text{H}_2\text{O}$; **d**), transjunctional voltage steps do not elicit junctional currents, indicating that the cells are not electrically coupled. Capacity transients in **c** and **d** have been truncated.

Table 1 Ionic coupling in paired oocytes expressing Shak-B

RNA injected (cell/cell)	Junctional conductance, g_j (μS)
Water/water – antisense Cx38	0.043 ± 0.026 (15)
Water/water + antisense Cx38	0.016 ± 0.006 (18)
L/L	15.87 ± 1.45 (30)
N/N	0.012 ± 0.006 (16)

Junctional conductance (g_j) is I_j/V_j (see Fig. 2), and was calculated, in all cases, at depolarizing voltage steps of 10 – 20 mV , well below the threshold necessary to induce voltage-sensitive reductions in I_j (see Fig. 2). So g_j is the mean ($\pm$ s.e.m.) of the maximum steady-state conductance for the number of cell pairs shown in brackets. Data are from oocyte pairs in which both cells had resting membrane potentials of -40 mV or greater. The low mean conductance in water/water pairs that were not pretreated with Cx38 antisense oligonucleotides (–antisense) was effectively eliminated by 24 h preinjection of antisense (+ antisense). L/L and N/N pairs were antisense pretreated. L, *shak-B(lethal)*; N, *shak-B(neural)*.

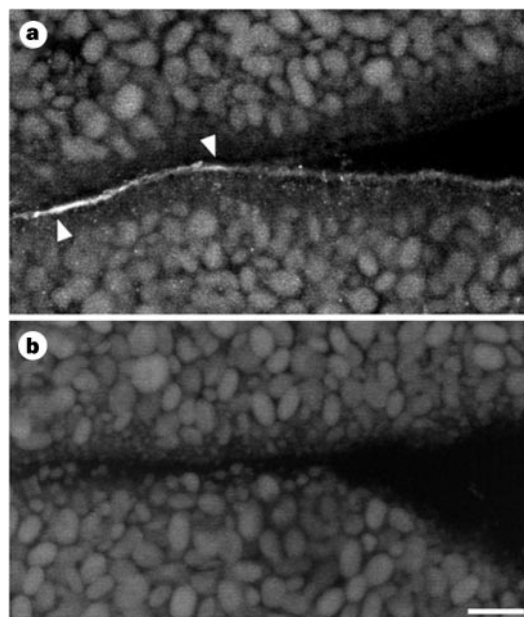


Figure 3 Shak-B(neural) protein is located in the region of cell–cell contact of RNA-injected paired oocytes. Oocyte pairs were sectioned and labelled with Shak-B antiserum. Confocal image of immunofluorescence at the site of plasma membrane apposition (arrowheads) of a pair 48 h after injection of 50 ng *shak-B(neural)* RNA (**a**) or water (**b**). Scale bar, $100\text{ }\mu\text{m}$.

from the presence of endogenous connexin38 (Cx38), which forms intercellular junctions with a low total conductance in some batches of oocytes^{14,15}.

Most water-injected oocyte pairs exhibited no measurable electrical coupling. Of 15 pairs recorded without prior injection of Cx38 antisense oligonucleotides (Table 1, water/water – antisense), only one showed significant coupling (junctional conductance, g_j , 0.39 μ S). We found no evidence of endogenous coupling in water/water pairs preinjected with oligonucleotides antisense to Cx38 (Fig. 2d, Table 1, + antisense). This procedure was used, therefore, to ensure that Cx38 channels did not contribute to junctional conductance in oocyte pairs expressing Shak-B.

In contrast to water/water controls, pairs in which both cells were injected with 0.05–0.5 ng of *shak-B(lethal)* RNA reliably formed intercellular channels (Fig. 2a, L/L) within 24 h. The average steady-state junctional conductance for pairs injected with 0.5 ng of RNA was 15.87 μ S (Table 1). Reducing the concentration of the RNA lowered the average junctional conductance (0.1 ng, $g_j = 7.35 \pm 2.38 \mu$ S, $n = 3$; 0.05 ng, $g_j = 3.66 \pm 1.17 \mu$ S, $n = 5$; 0.01 ng, $g_j = 1.49 \pm 0.64 \mu$ S, $n = 14$ including 5 pairs which showed no significant coupling). These values are within the range of junctional conductances reported for members of the connexin gene family expressed in *Xenopus* oocytes^{16,17}.

Because the oocytes were pretreated with Cx38 antisense oligonucleotides, intercellular channel formation in L/L pairs is presumably due solely to the translation of the exogenous RNA. As an additional control, we omitted the antisense pretreatment and paired *shak-B(lethal)* RNA with water-injected oocytes. This combination did not form channels, suggesting that the endogenous Cx, if present, is not competent to form heterotypic channels (each hemichannel composed of a different protein) with the *Drosophila* protein. We conclude, therefore, that expression of Shak-B(lethal) leads to electrical coupling through homotypic interactions.

To compare the electrical properties of Shak-B(lethal) junctions with connexin and previously characterized *Drosophila* gap junctions, we determined the voltage sensitivity of the channels. The junctional conductance in L/L oocyte pairs depends on the transjunctional voltage, that is, the voltage difference between the two cells, V_j (Fig. 2a). For $V_j \leq 40$ mV, the junctional current (I_j) is constant over time and proportional to V_j and hence g_j is constant. For $V_j \geq 60$ mV, the relationship between current and voltage is nonlinear; I_j decreases over time (with a time constant that is shorter for higher V_j) to a non-zero steady state. The near steady-state g_j (measured at the end of the step change; Fig. 2a) is therefore reduced. For the two pairs shown in Fig. 2a, g_j decreased from 10.7 to 7.04 μ S and from 5.33 to 2.83 μ S, respectively, when the voltage was stepped from –80 to 0 mV. The greater V_j sensitivity of cell pairs with lower maximal junctional conductance (compare the time course of inactivation of I_j in the L/L pairs in Fig. 2a) may be an artefact of the voltage-clamp procedure. Inadequate voltage control across the channels of larger gap junctions, which have higher cytoplasmic access resistance, is predicted to lead to overestimation of steady-state g_j and hence apparently reduced sensitivity to V_j (ref. 18). With this caveat in mind, Shak-B(lethal) junctions appear to be less sensitive to V_j than, for example, Cx38 junctions. For junctions with a similar maximal steady-state g_j , the V_j required to reduce g_j is lower, the time course of inactivation is faster, and the residual conductance is lower for Cx38 than for Shak-B(lethal) junctions (compare Fig. 3 in ref. 19 with L/L pair, lower panel in Fig. 2a).

The voltage-clamp protocol used in Fig. 2a does not distinguish transjunctional voltage from inside-outside voltage (V_{i-o}) imposed between the cytoplasm and the extracellular space. Figure 2b demonstrates, however, that L/L junctions ($g_j = 1–15 \mu$ S) are insensitive to V_{i-o} ; steady-state g_j is constant at membrane potentials between –60 mV and +40 mV. The partial closure of the junctions observed in Fig. 2a thus is due solely to changes in V_j . In this respect Shak-B(lethal) junctions differ from the gap junctions physiologi-

cally characterized in *Drosophila* salivary glands and embryonic muscle^{20,21}, which are gated by both V_j and V_{i-o} , suggesting that they are composed of channel proteins other than Shak-B(lethal).

Unlike *shak-B(lethal)*, the injection of *shak-B(neural)* RNA never induced electrical coupling (N/N, Fig. 2c); at 0.5 ng RNA the average junctional conductance was 0.012 μ S, which is not significantly different from the intercellular conductance of Cx38 antisense pretreated water/water pairs (Table 1). We tested higher concentrations of RNA, up to 50 ng, but similarly found no evidence of channel formation.

We are confident that the failure of *shak-B(neural)* RNA to induce the formation of functional intercellular channels in oocytes is not due to an error in the coding sequence. In *Drosophila*, expression of the *shak-B(neural)* cDNA used here under the control of a mesoderm-specific promoter restores dye coupling between muscle cells in *shak-B(neural)* null mutant embryos (R.A.B. *et al.*, unpublished data). An alternative explanation is that the translated protein, although clearly associated with the oocyte membrane fraction (Fig. 1), is not correctly targeted to the plasma membrane. We used immunocytochemistry to define further the distribution of Shak-B protein in oocyte pairs injected with *shak-B(neural)* RNA. The protein is concentrated at the region of apposition between the cells in or close to the plasma membranes (Fig. 3a). At least superficially, therefore, it appears to be ideally positioned to form intercellular channels. No corresponding labelling is seen in water-injected controls (Fig. 3b).

Like *shak-B(neural)*, three members of the connexin family fail to form homotypic channels in paired oocytes^{15,17,22}. One suggestion is that these proteins instead interact with other family members to assemble either heterotypic or heteromeric (individual hemichannels composed of two or more proteins) channels. In preliminary experiments we found that *shak-B(neural)*-injected oocytes did not form heterotypic channels with *shak-B(lethal)*-injected cells; it is possible, however, that the former assembles heterotypic or heteromeric channels with other members of the family which have been identified in *Drosophila*^{3,6} (J.A.D., unpublished data). Alternative possibilities are that the assembly or gating of Shak-B(neural) channels is regulated in a cell-specific manner, or that the protein itself is a regulatory rather than a structural component of gap junctions.

Although the precise function of Shak-B(neural) at gap junctions remains ambiguous, we have demonstrated that Shak-B(lethal) can form intercellular channels. This provides the first direct evidence for a family of gap junction proteins functionally analogous to, but molecularly distinct from, the connexins. It raises questions about the evolution of the molecular components of gap junctions. More importantly, it means that genetic approaches in *Drosophila* can now help to determine the roles of direct cell–cell communication in developing and mature systems. □

Methods

Preparation of RNA. DNA fragments containing the complete coding region of *shak-B(neural)* or *shak-B(lethal)* were subcloned into the plasmid vector pSPIC2L. Plasmids were linearized with *Xho*I and transcribed with SP6 RNA polymerase (Boehringer Mannheim) in the presence of m⁷G(5')ppp(5')G (Boehringer Mannheim) to produce capped RNAs. The integrity of the RNA and the effectiveness of capping were tested by translating 0.5–1 μ g *in vitro* in a rabbit reticulocyte lysate system (Promega, according to the manufacturer's instructions) and in a cytoplasm extract of *Xenopus* eggs²³.

Oocyte isolation, microinjection and pairing. The procedure is essentially as described previously^{16,24}. Defolliculated stage V–VI oocytes were preinjected with Cx38 DNA antisense oligonucleotides (sequence 5'-CTGACTGCTCGT-CTGTCCACACAG-3', which is antisense to nucleotides 204–227 in *Xenopus* Cx38; 20 ng (unmodified) in 20 nl RNAase-free water). One day later they were microinjected with 0.01–50 ng RNA in a volume of 20 nl RNAase-free water or water only (controls), the vitelline envelope was removed, and cells were paired and left for 24–48 h in Barth's saline at 19 °C.

Immunocytochemistry. Oocyte pairs were fixed in 4% paraformaldehyde in phosphate-buffered saline (PBS) and embedded in acrylamide²⁵. Cryostat sections (14–20 µm) were labelled with Shak-B antiserum (1:400 in blocking solution: 50 mM l-lysine, 4% normal goat serum, 0.2% Triton X-100 in PBS) followed by an FITC-conjugated secondary antibody (Pierce Warriner, 1:100 in blocking solution) and mounted in Citifluor (Agar Scientific). Sections were imaged with a confocal microscope (MRC 600, Bio-Rad).

Electrophoresis and western blotting. Oocyte membranes²⁴ and *Drosophila* nervous systems were solubilized in SDS gel-loading buffer, containing 8 M urea, at room temperature. Proteins were electrophoretically separated on 12.5% SDS–polyacrylamide gels and transferred, in CAPSO buffer (25-mM CAPSO (Sigma), 20% methanol, pH 10), to nitrocellulose (Hybond-C, Amersham). Blots were labelled with Shak-B antiserum (1:500 in blocking solution: 5% Marvel, 2% bovine serum albumin, 0.1% Tween-20 in PBS) followed by a peroxidase-conjugated secondary antibody (Sigma, 1:1000 in blocking solution) and developed with ECL (Amersham). Gels were scanned with a laser densitometer (Sharp) using Imagemaster ID (Pharmacia).

Electrophysiology. Individual oocytes of pairs were impaled with two borosilicate glass microelectrodes (1–4 mΩ) filled with a solution of 3 M KCl, 10 mM HEPES, 10 mM EGTA, pH 7.5, and recorded using the double voltage-clamp procedure²⁶. Protocols to determine junctional conductance (g_j) and its sensitivity to transjunctional voltage (V_j) and inside-outside voltage (V_{i-o}) were as described²⁰.

Received 18 August; accepted 16 September 1997.

1. Bruzzone, R., White, T. W. & Paul, D. L. Connections with connexins: the molecular basis of direct intercellular signalling. *Eur. J. Biochem.* **238**, 1–27 (1996).
2. Kumar, N. M. & Gilula, N. B. The gap junction communication channel. *Cell* **84**, 381–388 (1996).
3. Crompton, D., Todman, M., Wilkin, M., Ji, S. & Davies, J. Essential and neural transcripts from the *Drosophila* *shaking-B* locus are differentially expressed in the embryonic mesoderm and pupal nervous system. *Dev. Biol.* **170**, 142–158 (1995).
4. Krishnan, S. N., Frei, E., Swain, G. P. & Wyman, R. J. *Passover*: a gene required for synaptic connectivity in the giant fibre system of *Drosophila*. *Cell* **73**, 967–977 (1993).
5. Krishnan, S. N., Frei, E., Schalet, A. P. & Wyman, R. J. Molecular basis of intracystic complementation in the *Passover* locus of *Drosophila*. *Proc. Natl Acad. Sci. USA* **92**, 2021–2025 (1995).
6. Watanabe, T. & Kankel, D. R. Molecular cloning and analysis of *ll(1) oge*, a locus of *Drosophila melanogaster* with prominent effects on the postembryonic development of the central nervous system. *Genetics* **126**, 1033–1044 (1990).
7. Starich, T. A., Herman, R. K. & Shaw, J. E. Molecular and genetic analysis of *unc-7*, a *Caenorhabditis elegans* gene required for coordinated locomotion. *Genetics* **133**, 527–541 (1993).
8. Starich, T. A., Lee, R. Y. N., Panzarella, C., Avery, L. & Shaw, J. E. *eat-5* and *unc-7* represent a multigene family in *Caenorhabditis elegans* involved in cell–cell coupling. *J. Cell Biol.* **134**, 537–548 (1996).
9. Barnes, T. M. OPUS: a growing family of gap junction proteins? *Trends Genet.* **10**, 303–305 (1994).
10. Phelan, P. *et al.* Mutations in *shaking-B* prevent electrical synapse formation in the *Drosophila* giant fibre system. *J. Neurosci.* **16**, 1101–1113 (1996).
11. Homyk, T., Szidonya, J. & Suzuki, D. T. Behavioral mutants of *Drosophila melanogaster* III. Isolation and mapping of mutations by direct visual observations of behavioral phenotypes. *Mol. Gen. Genet.* **177**, 553–565 (1980).
12. Thomas, J. B. & Wyman, R. J. Mutations altering synaptic connectivity between identified neurons in *Drosophila*. *J. Neurosci.* **4**, 530–538 (1984).
13. Willecke, K. & Haubrich, S. Connexin expression systems: To what extent do they reflect the situation in the animal? *J. Bioenerg. Biomembr.* **28**, 319–326 (1996).
14. Barrio, L. C. *et al.* Gap junctions formed by connexins 26 and 32 alone and in combination are differently affected by applied voltage. *Proc. Natl Acad. Sci. USA* **88**, 8410–8414 (1991).
15. Hennemann, H. *et al.* Two gap junction genes, connexin 31.1 and 30.3, are closely linked on mouse chromosome 4 and preferentially expressed in skin. *J. Biol. Chem.* **267**, 17225–17233 (1992).
16. Swenson, K. I., Jordan, J. R., Beyer, E. C. & Paul, D. L. Formation of gap junctions by expression of connexins in *Xenopus* oocyte pairs. *Cell* **57**, 145–155 (1989).
17. Bruzzone, R., White, T. W. & Paul, D. L. Expression of chimeric connexins reveals new properties of the formation and gating behavior of gap junction channels. *J. Cell Sci.* **107**, 955–967 (1994).
18. Wilders, R. & Jongsma, H. J. Limitations of the dual voltage clamp method in assaying conductance and kinetics of gap junction channels. *Biophys. J.* **63**, 942–953 (1992).
19. Ebihara, L., Beyer, E. C., Swenson, K. I., Paul, D. L. & Goodenough, D. A. Cloning and expression of a *Xenopus* embryonic gap junction protein. *Science* **243**, 1194–1195 (1989).
20. Verselis, V. K., Bennett, M. V. L. & Bargiello, T. A. A voltage-dependent gap junction in *Drosophila melanogaster*. *Biophys. J.* **59**, 114–126 (1991).
21. Gho, M. Voltage-clamp analysis of gap junctions between embryonic muscles in *Drosophila*. *J. Physiol. (Lond.)* **481**, 371–383 (1994).
22. Bruzzone, R., White, T. W., Yoshizaki, G., Patino, R. & Paul, D. L. Intercellular channels in teleosts: functional characterization of two connexins from Atlantic croaker. *FEBS Lett.* **385**, 301–304 (1995).
23. Chevalier, S., Tassan, J.-P., Cox, R., Philippe, M. & Ford, C. Both *cdc2* and *cdc2* proteins promote S phase initiation in frog egg extracts. *J. Cell Sci.* **108**, 1831–1841 (1995).
24. Colman, A. in *Transcription and Translation—A Practical Approach* (ed. Hames, B. D. & Higgins, S. J.) 271–302 (IRL, Oxford, 1984).
25. Hausen, P. & Dreyer, C. The use of polyacrylamide as an embedding medium for immunohistochemical studies of embryonic tissues. *Stain Technol.* **56**, 287–293 (1981).
26. Spray, D. C., Harris, A. L. & Bennett, M. V. L. Equilibrium properties of a voltage-dependent junctional conductance. *J. Gen. Physiol.* **77**, 77–93 (1981).

Acknowledgements. We thank H. Woodland for the pSP/C2L vector, and E. Jones, S. King, M. O'Shea, D. Paul, M. Stern, M. Todman and J. Yeoman for help, advice and discussion. This work was funded by the BBSRC, UK.

Correspondence and requests for materials should be addressed to P.P. (e-mail: bafb5@central.sussex.ac.uk).

A serine/threonine kinase gene defective in Peutz–Jeghers syndrome

Akseli Hemminki*, David Markie†, Ian Tomlinson‡§, Egle Avizienyte*, Stina Roth*, Anu Loukola*, Graham Bignell‡, William Warren‡, Maria Aminoff||, Pia Höglund*, Heikki Järvinen†, Paula Kristo*, Katarina Pelin*, Maaret Ridanpää*, Reijo Salovaara*, Tumi Toro†, Walter Bodmer*, Sylviane Olschwang**, Anne S. Olsen††, Michael R. Stratton‡, Albert de la Chapelle*** & Lauri A. Aaltonen*

Departments of * Medical Genetics and † Pathology, Haartman Institute, PO Box 21 (Haartmaninkatu 3), 00014 University of Helsinki, Finland

‡ Department of Pathology, Dunedin School of Medicine, PO Box 913, Dunedin, New Zealand

§ Section of Cancer Genetics, Haddow Laboratories, Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey SM2 5NG, UK

§ Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Oxford OX3 9DU, UK

¶ Second Department of Surgery, Helsinki University Central Hospital, Haartmaninkatu 4, 00290 Helsinki, Finland

|| Folkhälsan Institute of Genetics, Mannerheimintie 97, 00280 Helsinki, Finland

* ICRF Cancer and Immunogenetics Laboratory, Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DS, UK

** Inserm U434, Fondation Jean Dausset-CEPH F75010 Paris, France

†† Human Genome Center, L-452, Lawrence Livermore National Laboratory, Livermore, California 94550, USA

*** Comprehensive Cancer Center, Ohio State University, 420 W 12th Avenue, Columbus, Ohio 43210, USA

Studies of hereditary cancer syndromes have contributed greatly to our understanding of molecular events involved in tumorigenesis. Here we investigate the molecular background of the Peutz–Jeghers syndrome^{1,2} (PJS), a rare hereditary disease in which there is predisposition to benign and malignant tumours of many organ systems. A locus for this condition was recently assigned to chromosome 19p (ref. 3). We have identified truncating germline mutations in a gene residing on chromosome 19p in multiple individuals affected by PJS. This previously identified but unmapped gene, *LKB1* (ref. 4), has strong homology to a cytoplasmic *Xenopus* serine/threonine protein kinase *XEEK1* (ref. 5), and weaker similarity to many other protein kinases. Peutz–Jeghers syndrome is therefore the first cancer-susceptibility syndrome to be identified that is due to inactivating mutations in a protein kinase.

Several genes are associated with a hereditary predisposition to cancer. In many cases, for example familial retinoblastoma, hereditary cancer syndromes display a limited tumour spectrum characteristic of the condition. This phenomenon may be related to the gatekeeper function of the predisposing genes in the respective tissues. Gatekeeper genes for many important organ systems remain to be identified⁶. Peutz–Jeghers syndrome (PJS)^{1,2} is a rare, dominantly inherited condition characterized primarily by predisposition to benign intestinal hamartomatous polyps. Carriers of the gene are also predisposed to many classes of cancer, including cancers of the gastrointestinal tract, breast, testis and ovary^{7,8}. In addition to tumour susceptibility, the syndrome is characterized by mucocutaneous pigmentation which usually affects the lips, buccal mucosa and digits.

A locus for PJS has been mapped to chromosome 19p (ref. 3). Initial analysis of multiple hamartomas derived from a single PJS patient (SL8) by comparative genomic hybridization revealed deletions of chromosome 19p. Further evaluation using polymorphic microsatellite markers in this region indicated that the deletions

were exclusively on the chromosome inherited from the unaffected parent, raising the possibility that the target of the deletions was a tumour-suppressor gene underlying PJS³. Linkage analysis in PJS families using chromosome 19p markers confirmed the presence of a high-penetrance susceptibility locus telomeric to marker D19S565.

A search for the predisposing gene was made by using different

approaches; efforts were facilitated by a complete cosmid contig across the putative PJS gene region, created as part of the human chromosome-19 mapping project (ref. 9, and <http://www-bio.llnl.gov/>). The PJS region was first narrowed down to 800 kilobases (kb) around markers D19S886 and D19S883 by meiotic recombination mapping in PJS families using microsatellite and biallelic polymorphisms developed for this purpose. Transcripts in this interval were identified by database searches and by solution hybridization of complementary DNA derived from human testis or colon RNA to genomic DNA from cosmids mapping to the PJS gene region (direct cDNA selection^{10,11}). Twenty-seven transcripts were screened for mutations, typically by sequencing of products obtained by reverse transcription followed by polymerase chain reaction (RT-PCR) from lymphoblastoid cell lines derived from PJS patients. In some cases, Southern blot analysis of genomic DNA was also done. Among the transcripts detected by cDNA selection was *LKB1*, a gene predicted to encode a 433-amino-acid protein showing strong sequence similarity to the serine/threonine protein kinases (GenBank accession number U63333; ref. 4). The genomic location of *LKB1* has not previously been reported, but, using genomic cosmid sequence data, we confirmed that *LKB1* is located between the genes *CIRP* and *GPX4*, in accordance with the cDNA selection results (a map of genes in this area is available at <http://www-bio.llnl.gov/>).

Primers were designed to analyse mutations in the *LKB1* cDNA. The sequence available in the database covered only the translated region, and hence it was not possible to amplify the whole coding region for mutational screening purposes. However, the primers chosen amplified most of the coding sequence (87%), and this could be screened for mutations by RT-PCR and sequencing in 12 familial PJS patients. Each of the 12 families was compatible with linkage of the PJS phenotype to chromosome 19p. Evidence of a putative mutation in this gene was first obtained from the case (SL8) whose polyps were originally used in the localization of PJS. RT-PCR products from this individual gave a smaller cDNA fragment in addition to the normally sized band. Sequence analysis revealed a heterozygous 188-bp deletion (nucleotides 921–1,108) in the cDNA which is predicted to generate a translational frameshift (Table 1). Subsequently, eight further heterozygous mutations predicted to cause truncation of the encoded protein were identified among the

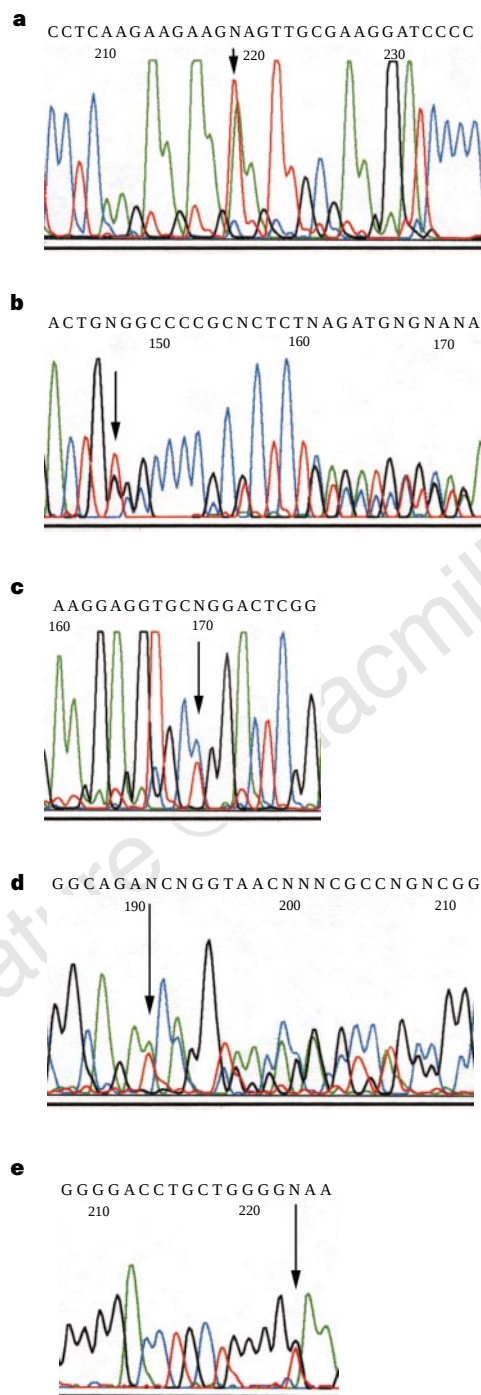


Figure 1 Examples of heterozygous mutations identified in *LKB1* gene in PJS patients. Genomic sequences (see Methods) are presented in the 5' to 3' direction and the location of the mutation is marked by an arrow. **a**, Patient SL32: a nonsense mutation in codon 84 (lysine → stop). **b**, Patient SL31: a 2-bp deletion in codons 277–278. **c**, Patient SL25: a missense mutation in codon 67 (leucine → proline). **d**, Patient SL26: a 9-bp deletion in codons 303–306. **e**, Patient SL12: a nonsense mutation in codon 57 (glutamic acid → stop).

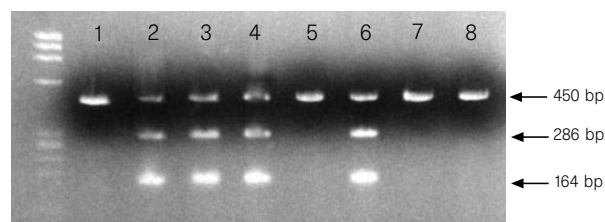
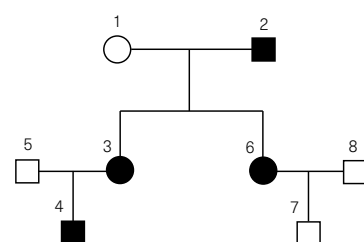


Figure 2 Segregation of a nonsense mutation in a PJS family including patient SL20 (lane 3). The heterozygous mutation is detected by *AccI* digestion of a 450-bp PCR product. The mutant allele (displayed as two fragments of sizes 286 and 164 bp) segregates with the disease phenotype. The numbers above the lines correspond with those in the pedigree. A molecular size marker (Φ X174) is shown on the left.

Table 1 *LKB1* sequencing results in the 12 PJS patients

Patient	Codon	Nucleotide change	Predicted effect	Control individuals
SL8	307-370	188-bp deletion	Frameshift, stop at codon 404	0/86
SL12*	57	G → T	Glutamic acid → stop	0/98
SL14*†	66-75	29-bp deletion	Frameshift, stop at codon 152	0/98
SL20*†	70	G → T	Glutamic acid → stop	0/98
SL25*†	67	T → C	Leucine → proline	0/98
SL26*†	303-306	9-bp deletion	Ile Arg Gln His → Asn	0/49
SL27*†	60	C → G	Tyrosine → stop	0/98
SL28*	55-57	1-bp insertion	Frameshift, stop at codon 162	0/98
SL29	98-155	174-bp deletion	Truncated (in-frame) product of 376 amino acids	0/59
SL30		No changes		
SL31*†	277-278	2-bp deletion	Frameshift, stop at codon 283	0/72
SL32*†	84	A → T	Lysine → stop	0/66

* Confirmed by genomic sequencing.

† Variant shown to segregate with the PJS phenotype in the respective family.

1	<i>LKB1</i>	-----ME	VVDPQQLGMFTEGEL	MSVGMOTFIHRIDST	32
2	<i>XEEK1</i>	-----MLCPS	SMDEEGSEETGFLGD	LSVGMOTFIHRIDST	35
3	<i>mLKB1</i>	WRSRRXEDKEWARM	VADPEPLGLFSEGEL	MSVGMOTFIHRIDST	45
1	<i>LKB1</i>	EVIIYQPRRKRAKLIG	KYLMGDLLEGESYGK	VKEVLDSSETLCRRAY	77
2	<i>XEEK1</i>	EVIIYQPRRKRAKLIG	KYLMGDLLEGESYGK	VKEMLDSDTLCRRAY	80
3	<i>mLKB1</i>	EVIIYQPRRKRAKLIG	KYLMGDLLEGESYGK	VKEVLDSSETLCRRRS	96
1	<i>LKB1</i>	KILKKKKLRRIPNGE	ANVKKEIQLLRLRH	KNVIQLVDVLYNEEK	120
2	<i>XEEK1</i>	KILKKKKLRRIPNGE	ANVKKEIQLLRLRH	RNVIQLVDVLYNEEK	120
3	<i>mLKB1</i>	KILKKKKLRRIPNGE	ANVKKEIQLLRLRH	-----	120
1	<i>LKB1</i>	QKMYMVEYCVCGMQ	EMLDSVPEKRFVPCQ	AHGYFCQLIDGLEYL	167
2	<i>XEEK1</i>	QKMYMVEYCVCGMQ	EMLDSVQDKHFPVFQ	AHGYFCQLIDGLEYL	170
3	<i>mLKB1</i>	-----	-----	-----	120
1	<i>LKB1</i>	HSQGIYHKDIKPGNL	LLTTGGTLKISDLGV	AEALHPFAADDTCT	212
2	<i>XEEK1</i>	HSQGIYHKDIKPGNL	LLTTGGTLKISDLGV	AEALHPFAEGDTCRT	215
3	<i>mLKB1</i>	-----	-----	-----	120
1	<i>LKB1</i>	SQGSAPAFQPEIANG	LDTFSGFVKDIWSAG	VTLYNITTGLYPFEG	257
2	<i>XEEK1</i>	SQGSAPAFQPEIANG	LDTFSGFVKDIWSAG	VTLYNITTGLYPFEG	260
3	<i>mLKB1</i>	-----	-----	-----	120
1	<i>LKB1</i>	DNIYKLFENIGKGSY	AIPGDCGPPSLDLLK	GMLEYEPAKRFSIRQ	307
2	<i>XEEK1</i>	DNIYKLFENIGKGSY	SIPEECGPPSLDLLR	GMLEYEPAKRFSIRQ	307
3	<i>mLKB1</i>	-----	-----	-----	120
1	<i>LKB1</i>	IRQHSWFRKKHPPAE	APVPIPPSPDTKDRW	RSMTVVPYLEDLHGA	347
2	<i>XEEK1</i>	IRQHNWFRKKHPPMD	PVPIPPSPETKDRW	RSLTVVPYLEDLHGY	350
3	<i>mLKB1</i>	-----	-----	-----	120
1	<i>LKB1</i>	DEDEDLFDIEDDIY	TQDFTVPGQVPEEEA	SHNGQRRGLPKAVCM	397
2	<i>XEEK1</i>	SEEDLDFEDDIY	TQDFTVPGQVAEDDY	FAQTQSTAPSKQLCM	397
3	<i>mLKB1</i>	-----	-----	-----	120
1	<i>LKB1</i>	NGTEAAQLSTKSRAE	GRAPNPARKACASS	KIRRLSACKQQ	437
2	<i>XEEK1</i>	NGTES-QLKTERRVS	---SSSQKASTTGS	KVRKLSACKQQ	437
3	<i>mLKB1</i>	-----	-----	-----	120

Figure 3 Sequence alignment of *LKB1* and *XEEK1* proteins and a homologous mouse EST (*mLKB1*) predicted protein sequence created by using the ClustalW1.7 e-mail server. Numbers refer to amino acids in the respective sequences. The protein kinase domain presenting a strong consensus sequence locates between codons 50 and 250 in the human *LKB1* sequence. The analysis of the mouse EST reading frames revealed a possible sequencing error at the nucleotide position 265 (ref. 13). A putative loss of a single nucleotide G in the GenBank sequence at this site has led to a frameshift and loss of homology for the rest of the predicted mouse amino-acid sequence. The homology between human and mouse protein sequence was retained (as shown here) by adding a nucleotide G at position 265 in the mouse sequence.

12 patients (Table 1, Fig. 1). The nature of the mutations is consistent with the idea that the gene is inactivated during oncogenesis and conforms with the model of a recessive tumour-suppressor gene suggested by allele loss in PJS polyps³. Two non-truncating changes were detected, one an amino-acid substitution,

the other a deletion of 9 bp. Although these were not detected in 98 and 49 control individuals respectively (Table 1), at present we cannot exclude the possibility that they represent rare polymorphisms. Owing to an ongoing large-scale genomic sequencing effort in the Lawrence Livermore National Laboratory (<http://www-bio.llnl.gov>), we were able to design genomic primers for detection of these two changes, as well as for seven of the other changes, enabling us to confirm the genomic presence of these nine sequence variants and to study segregation of seven of the variants in the respective PJS families. All seven changes segregated with the PJS phenotype (Table 1 and Fig. 2). One of the 12 PJS cases screened does not currently show sequence variants in *LKB1*. The absence of mutations in this case may be due to incomplete screening of the gene, insensitivity of our mutation-detection methods, transcript instability as a result of the presence of a mutation, large genomic deletions that are undetectable by the PCR-based approach used, or genetic heterogeneity, with mutations present at another distinct locus. Northern blotting and RT-PCR analysis indicated that *LKB1* is expressed in all tissues examined: liver, skeletal muscle, kidney, pancreas, heart, brain, placenta, lung, spleen, thymus, testis, peripheral blood leukocyte, colon, ovary, small intestine and prostate.

LKB1 was originally identified as a serine/threonine protein kinase expressed in human fetal liver and shows weak sequence homology to many protein kinases over the conserved catalytic core of the kinase domain (between codons 50–337) that is common to both serine/threonine and tyrosine protein kinase family members. A mouse homologue has been reported that shows 88% nucleotide sequence identity over a 320-bp overlap with the human sequence and which was identified through a systematic search of genes expressed in embryo implantation sites¹². Further sequence database searches revealed that *LKB1* is also highly homologous to the *Xenopus* serine/threonine kinase *XEEK1* (for *Xenopus* egg and embryo kinase)⁵ to which it has 83.7% amino-acid sequence identity over 416 codons (Fig. 3) (<http://www.ncbi.nlm.nih.gov/BLAST/> and <http://www2.ebi.ac.uk/fasta3>). *XEEK1* is a 432-amino-acid cytosolic protein with narrow substrate specificity and which may itself be phosphorylated by protein kinase A⁵. It is expressed in *Xenopus* oocytes and fertilized eggs but messenger RNA levels decline after fertilization and are barely detectable during gastrulation and at later embryonic stages.

The results indicate that *LKB1* is the main gene that when mutated is responsible for familial PJS. The predicted protein sequence suggests that *LKB1* encodes a protein kinase which has been confirmed in studies of the likely *Xenopus* homologue⁵. The mutations found in the gene are likely to abolish some or all of the kinase activity of the protein. Although activation of kinase activity may be responsible for cancer susceptibility in multiple endocrine neoplasia type 2 (*RET*)¹³, familial renal papillary cancer (*MET*)¹⁴ and in familial melanoma (*CDK4*)¹⁵, to our knowledge this is the first gene that predisposes to cancer as a result of disabling its encoded kinase activity.

No major disturbances of development are seen in PJS individuals heterozygous for the *LKB1* defect. It is likely that initiation of hamartoma formation occurs after somatic inactivation of the wild-type *LKB1* allele³. The function of *LKB1* in human development remains obscure, but may be revealed by model systems, such as mice homozygous for the defect.

It should now be possible to offer predictive genetic testing to members of PJS families. The relatively small size of the gene and the nature of the mutations should facilitate the implementation of such diagnostic procedures. □

Methods

Patient samples and RNA. Cell lines were prepared from PJS patients' blood samples using standard methods, and RNA was extracted using the RNeasy kit (Qiagen).

RT-PCR. 20 µl cDNA was created from 0.8 µg of each of the patients' RNA samples using standard random priming methods with M-MLV reverse transcriptase (Promega) and RNase inhibitor (Promega). RT-PCR was usually performed in DNA Engine thermal cyclers (MJ Research) under the following conditions: 3 µl cDNA, 1 × buffer (with MgCl₂), 250 µM dNTPs, 0.8 µM reverse and forward primers, 2 units of Amplitaq Gold polymerase (Perkin Elmer) and water in a final volume of 50 µl. Primers were designed for each transcript studied using primer 3 (<http://www-genome.wi.mit.edu/>). For *LKB1*, only the protein-coding sequence without surrounding untranslated regions was available. In most cases, the following primer pair was used to amplify 1,135 of the total 1,302 bp (87%) of the reading frame (each primer 5' to 3'): LKBF1: GAGCTGATGTCGGTGGGTAT; LKBR5: GCCCTGGATTGGTGCTC. In addition to LKBF1 and LKBR5, the following internal primers were used for sequencing of the PCR product: LKBR6, GCTATGCAGTACTCCAGGC; LKBF2, ATGGAGTACTGCGTGTGTGG; LKBR2, GTTGTAGAGGGTGACCCAG; LKBF6, GGCCTGGACACCTTCTCC. For *LKB1*, the following PCR programme was used: 12 min at 95 °C, 40 cycles of: 95 °C for 45 s, 57 °C for 1 min, 72 °C for 2 min; final extension 10 min at 72 °C.

Sequencing of PCR products. Sequencing was performed using either ABI PRISM Dye Terminator or ABI PRISM dRhodamine cycle sequencing kits (Perkin-Elmer) according to the manufacturer's instructions, and reactions were run on an ABI 373 A or 377 sequencer (Perkin Elmer), respectively.

Direct cDNA selection. Selection of transcripts from the candidate region was made using solution hybrid capture according to ref. 11. Individual cosmid clones spanning the critical region were digested with *RsaI* alone and with *HaeIII* plus *RsaI*. These digests were combined and ligated to adapters produced by annealing the following oligonucleotides (a) CCTCTGAAGGTTCCAGAATGATAGTAATTCGT, and (b) CACGAATTC ACTATCGATTCTGGAA CCTTCAGAGGTTT. Selector DNA was then generated by PCR, using a 5'-biotinylated version of oligo(a). cDNA for capture was synthesized from testis poly(A)⁺ RNA primed by both oligo(dT) and random primers. Adapters formed from the preannealed oligonucleotides: AGCAAGTTCAGCCCTGGTTAAGGTCGACGTG and PO₄⁻ CACGTCGACC TTAACCAGGCTGAACCTTGC TTTT were ligated to cDNA, which was then amplified in the PCR using the primer AGCAAGTTCAGCCTGGTTAAG. After purification, selector DNA and cDNA were denatured and prehybridized with COT1 DNA (Gibco/BRL) to block repetitive sequences. Prehybridized selector DNA was then combined with prehybridized cDNA and the hybridization continued overnight. Hybrid DNA was captured from the hybridization mixture using streptavidin-coated magnetic beads. Trapped cDNA was then eluted and amplified by PCR to produce cDNA for a second round of hybrid capture. Amplified, secondary-selected cDNA was cloned into plasmid vectors and sequenced.

LKB1 expression. *LKB1* expression patterns were tested in a variety of human tissues. Northern blot hybridization was done using human multiple tissue northern blot filters I and II (Clontech) at 42 °C. The probe was amplified with primers LKBF1 and LKBR5 (see above) from cDNA. Radioactive labelling of the PCR product was performed with the rediprime DNA labelling system (Amersham) and the probe purified with NICK spin columns (Pharmacia Biotech). RT-PCR based analysis was performed with *LKB1* specific primers using human multiple tissue cDNA (MTC) panels I and II (Clontech) as a template.

Genomic PCR. The following intronic primers were used for genomic *LKB1* PCR amplifications: for nucleotides 1–290: GGAAGTCGGAACACAAGGAA and GGGAGGAGAGAAGGAAGGAA; for nucleotides 735–862: TCAACC ACCTTGACTGACCA and ACACCCCCAACCTACATTT; for nucleotides 863–920: GGAGTGGAGTGGCCTCTGT and CTCACCAGCTGCCACAT.

Received 8 October; accepted 28 November 1997.

1. Peutz, J. L. On a very remarkable case of familial polyposis of the mucous membrane of the intestinal tract and nasopharynx accompanied by peculiar pigmentation of the skin and mucous membrane. *Ned. Tijdschr. Geneesk.* **10**, 134–146 (1921).
2. Jeghers, H., McKusick, V. A. & Katz, K. H. Generalized intestinal polyposis and melanin spots of the oral mucosa, lips and digits: a syndrome of diagnostic significance. *N. Engl. J. Med.* **241**, 992–1005 (1949).
3. Hemminki, A. *et al.* Localization of a susceptibility locus for Peutz–Jeghers syndrome to 19p using comparative genomic hybridization and targeted linkage analysis. *Nature Genet.* **15**, 87–90 (1997).
4. Nezu, J. Molecular cloning of a novel serine/threonine protein kinase expressed in human fetal liver (unpublished). http://www.ncbi.nlm.nih.gov/ix/cgi-bin/birx_doc?genbank+41522.
5. Su, J.-Y., Erikson, E. & Maller, J. L. Cloning and characterization of a novel serine/threonine protein kinase expressed in early *Xenopus* embryos. *J. Biol. Chem.* **271**, 14430–14437 (1996).
6. Kinzler, K. W. & Vogelstein, B. Lessons from hereditary colorectal cancer. *Cell* **87**, 159–170 (1996).
7. Phillips, R. K. S., Spigelman, A. D. & Thomson, J. P. S. (eds) *Familial Adenomatous Polyposis and other Polyposis Syndromes* (Arnold, London, 1994).
8. Giardiello, F. M. *et al.* Increased risk of cancer in the Peutz–Jeghers syndrome. *N. Engl. J. Med.* **316**, 1511–1514 (1987).
9. Ashworth, L. K. *et al.* An integrated metric physical map of human chromosome 19. *Nature Genet.* **11**, 422–427 (1995).
10. Tagle, D. A., Swaroop, M., Lovett, M. & Collins, F. C. Magnetic bead capture of expressed sequences encoded within large genomic segments. *Nature* **361**, 751–753 (1993).
11. Futreal, P. A. *et al.* Isolation of a diverged homeobox gene, *MOX1*, from the *BRCA1* region on 17q21 by solution hybrid capture. *Hum. Mol. Genet.* **3**, 1359–1364 (1994).
12. Ko, M. S. H. *et al.* Systematic analyses of mouse genes expressed in embryo implantation site http://www.ncbi.nlm.nih.gov/ix/cgi-bin/birx_doc?genbank+431188.
13. Mulligan, L. M. *et al.* Germ-line mutations of the *RET* proto-oncogene in multiple endocrine neoplasia type 2A. *Nature* **363**, 458–460 (1993).
14. Schmidt, L. *et al.* Germline and somatic mutations in the tyrosine kinase domain of the *MET* proto-oncogene in papillary renal carcinoma. *Nature Genet.* **16**, 68–73 (1997).
15. Zuo, L. *et al.* Germline mutations in the p16(INK4a) binding domain of CDK4 in familial melanoma. *Nature Genet.* **12**, 97–99 (1996).

Acknowledgements. This study was aided by grants from the Helsinki University Central Hospital, the Academy of Finland, the Finnish Cancer Society, the Finnish Medical Society Duodecim, the Ida Montin Foundation, the Jalmari and Rauha Ahokas Foundation, the 350th Anniversary Foundation of the University of Helsinki, Leiras, the Emil Aaltonen Foundation, the UK Cancer Research Campaign, the ICRF, the Institute of Cancer Research, the U.S. Department of Energy, the Otago Medical School Bequest Funds, the Cancer Society of New Zealand and The Wellington Division of the Cancer Society of New Zealand. We thank K. Honkanen, M. Veini, T. Kosonen, S. Lindh, S. Lindroos, K. Collin, A. Georgescu, L. Gordon and R. Hamoudi for technical assistance; W. El-Rifai, K. Virtaneva, S. Ranta and S. Ezer for help with methodology; A. Moisio for control samples; D. Jenne for unpublished mapping data; the many PJS families for their collaboration; and the following clinicians for PJS samples: I. Ellis, S. Hodgson, T. Iwama, S. Loff, P. Zaubner, C. Marchere, J. Sampson, S. Davies, I. Talbot, J. Wyke, R. Houlston, O. Suomalainen, F. M. Giardiello, S. R. Hamilton, S. B. Gruber, C. Eng, S. A. Lynch, A. Hunter and G. Mettler, and the polyposis registry, St Mark's Hospital.

Correspondence and requests for materials should be addressed to L.A.A. (e-mail: Lauri.Aaltonen@Helsinki.Fi).

COPII–cargo interactions direct protein sorting into ER-derived transport vesicles

Meta J. Kuehn, Johannes M. Herrmann & Randy Schekman

Department of Molecular and Cell Biology, Barker Hall, Howard Hughes Medical Institute, University of California, Berkeley, California 94720, USA

Vesicles coated with coat protein complex II (COPII) selectively transport molecules (cargo) and vesicle fusion proteins from the endoplasmic reticulum (ER) to the Golgi complex^{1–3}. We have investigated the role of coat proteins in cargo selection and recruitment. We isolated integral membrane and soluble cargo proteins destined for transport from the ER in complexes formed in the presence of Sar1 and Sec23/24, a subset of the COPII components, and GTP or GMP-PNP. Vesicle fusion proteins of the vSNARE family and Emp24, a member of a putative cargo carrier family⁴, were also found in COPII complexes. The inclusion of amino-acid permease molecules into the complex depended on the presence of Shr3, a protein required for the permease to leave

the ER^{3,5}. Resident ER proteins Sec61, BiP (Kar2) and Shr3 were not included in the complexes, indicating that the COPII components bound specifically to vesicle cargo. COPII–cargo complexes and putative cargo adaptor–cargo complexes were also isolated from COPII vesicles. Our results indicate that cargo packaging signals and soluble cargo adaptors are recognized by a recruitment complex comprising Sar1–GTP and Sec23/24.

In *Saccharomyces cerevisiae*, the COPII components Sar1, Sec23/24 complex and Sec13/31 complex bind sequentially to the ER (S. Bednarek, T. Yeung and R.S., unpublished data), resulting in the formation of a transport vesicle². COPII vesicle budding and uncoating requires a cycle of GTP hydrolysis^{6,7}. Amino-acid permeases (Gap1 and Hip1)³, vSNAREs (Bet1, Bos1 and Sec22, which are integral membrane proteins that are part of the vesicle fusion machinery)⁸, and glycosylated pro- α factor (gp α F)¹ have been used to analyse the capture of integral membrane and soluble cargo molecules into COPII transport vesicles. Shr3, an integral membrane ER resident protein, is required for the permeases to be included into COPII vesicles both *in vitro* and *in vivo*^{3,5}.

We postulated that if coat proteins are involved in the selection of molecules for transport, then at some point before the emergence of a bud, cargo molecules and vesicle proteins might be bound directly or indirectly to the COPII components. Proteins in the p24 family, such as Emp24, have been implicated as cargo receptor/adaptor proteins that could serve as a link between luminal cargo molecules and coat proteins^{4,9–11}. Perhaps because of their redundancy, no severe secretion defects have been observed for null alleles of any of the eight yeast p24 genes (J.M.H., A. Rowley and R.S., unpublished data).

Cargo molecules in complex with coat proteins were isolated using a Sar1–glutathione S-transferase (GST) fusion protein. Sar1–GST supported the packaging of gp α F and amino-acid permeases into COPII vesicles, but required the nucleotide GMP–PNP rather than GTP, probably because Sar1–GST was a more potent GTPase than Sar1 (data not shown). Because GTP hydrolysis is not required for cargo sorting or vesicle budding², this restriction did not affect our assays. GDP- β S, an analogue of GDP that cannot be efficiently

phosphorylated and does not support cargo packaging from membranes¹², was used as a control in these studies.

Microsomes prepared from cells overproducing an epitope-tagged form of the histidine permease, Hip1–Myc, were incubated with some or all of the COPII components and Sar1–GST. Proteins bound to Sar1–GST were recovered from digitonin-solubilized membranes on glutathione–agarose (GSHA). The integral membrane proteins Hip1, Emp24, Sec22 and Bet1 were recovered on GSHA, presumably in a complex with Sar1–GST (Figs 1a, b, 2a). Incubations of the membranes with both Sar1–GST and Sec23/24 resulted in an average of sixfold greater recovery of membrane cargo compared with incubations using Sar1–GST alone. In incubations that included Sec13/31, little Sar1–GST–cargo complex was recovered from the ER membrane (Fig. 1a); instead, cargo was found in the vesicle-containing supernatant (data not shown), suggesting that the Sar1/Sec23/24 complexes were consumed during vesicle formation. Complexes were not recovered from membranes incubated with GST or in reactions containing GDP- β S (Figs 1b, 2a). Sec61, a resident ER membrane protein, does not enter COPII vesicles¹, and was not recovered on GSHA (Fig. 1a). These results demonstrate that integral membrane cargo molecules form specific complexes with activated Sar1 and Sec23/24 coat components before packaging into vesicles.

We next investigated whether soluble cargo, ³⁵S-radiolabelled gp α F, could be isolated in a complex with COPII components. Similar to integral membrane cargo, gp α F bound to GSHA in complexes isolated from membranes that had been incubated with Sar1–GST and Sec23/24 (Fig. 1b). Approximately 20-fold less gp α F was isolated on GSHA from reactions using Sar1–GST alone or using GDP- β S (Fig. 1b; Table 1, top). The soluble resident ER chaperone, Kar2 (BiP), which is not packaged into transport vesicles¹, did not form a complex with the COPII components under any reaction conditions (Fig. 1b). These results suggest that soluble ER cargo interacts specifically with COPII components, and imply that membrane-spanning adaptor proteins coordinate luminal cargo packaging by cytosolic coat proteins.

We tested whether Hip1–Myc and an epitope-tagged form of the

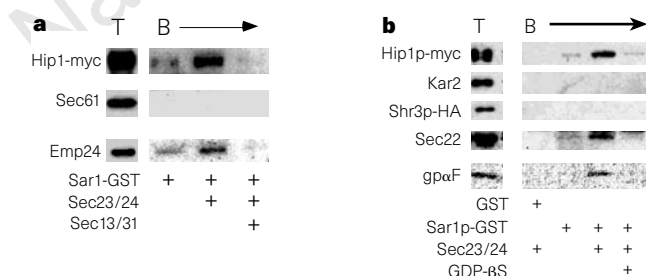


Figure 1 Vesicle proteins Hip1, Emp24, Sec22 and gp α F, but not ER resident proteins Sec61, Shr3 and Kar2, are present in complex with Sar1 and Sec23/24. **a**, Immunoblot analysis of 2- μ g total solubilized membranes (T) and complexes isolated from 300- μ l binding reactions (B) containing 100 μ g microsomes from PLY198[pPL321] (Hip1–Myc), GMP–PNP, 1 $\times$ ATP, and COPII components as indicated. **b**, Immunoblot analysis of 6- μ g total solubilized membranes (T) and complexes isolated from 300- μ l binding reactions (B) containing 225 μ g microsomes from PLY204[pPL327][pPL261] (Shr3–HA, Hip1–Myc), COPII components, and GMP–PNP or GDP- β S as indicated. gp α F: Phosphorimage of 3 μ g total solubilized membranes (T) and complexes isolated from 300- μ l binding reactions (B) containing 160 μ g ³⁵S-gp α F-loaded microsomes from PLY198, 1 $\times$ ATP, COPII components, and GMP–PNP or GDP- β S as indicated.

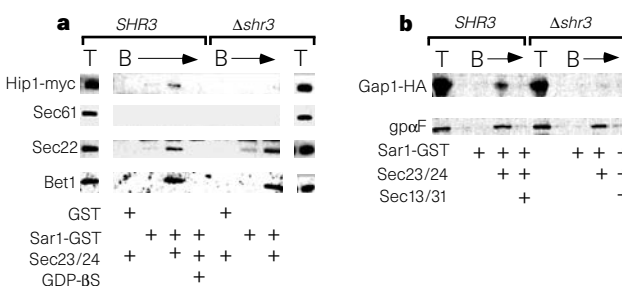


Figure 2 Shr3 is required specifically for incorporation of Hip1 and Gap1 into a complex with Sar1 and Sec23/24. **a**, Immunoblot of 4- μ g total solubilized membranes (T) and complexes isolated from 300- μ l binding reactions (B) containing 400 μ g microsomes from PLY214[pPL327] (SHR3, Hip1–Myc) or PLY229[pPL327] (Δ shr3, Hip1–Myc), COPII components, and GMP–PNP or GDP- β S as indicated. **b**, Immunoprecipitation of ³⁵S-Gap1–HA from 1.6 μ g total solubilized membranes (T) and complexes isolated from 300- μ l binding reactions (B) containing 40 μ g perforated spheroplasts from 3-min pulse-labelled PLY129[pPL269] (SHR3, Gap1–HA) or PLY134[pPL269] (Δ shr3, Gap1–HA). gp α F: Phosphorimage of ³⁵S-gp α F from 0.6- μ g total solubilized membranes (T) and complexes isolated from 300- μ l binding reactions (B) containing 25 μ g ³⁵S-gp α F-loaded, perforated spheroplasts from PLY129[pPL269] (SHR3, Gap1–HA) or PLY134[pPL269] (Δ shr3, Gap1–HA). Binding reactions contained 1 $\times$ ATP, and GMP–PNP, and COPII components as indicated. Quantitative analysis for **b** is shown in Table 1.

general amino-acid permease Gap1–haemagglutinin (HA)^{3,5} could form complexes with the COPII components in $\Delta shr3$ membranes. We found that Hip1–Myc recovery on GSHA was decreased using $\Delta shr3$ membranes (Fig. 2a). To quantify the permeases in COPII complexes, pulse-radiolabelled *SHR3* and $\Delta shr3$ cells were converted to perforated spheroplasts, incubated with the COPII proteins and Sar1–GST and solubilized. GSHA-bound radiolabelled Gap1–HA was immunoprecipitated and evaluated quantitatively. Approximately 1% of the radiolabelled Gap1–HA present in the ER membrane was recovered from incubations including Sar1–GST and Sec23/24 (Fig. 2b; Table 1, top). The low recovery of permease bound to GSHA may reflect the transient or labile nature of the complexes. Sixfold less COPII–Gap1–HA complex was recovered from reactions using $\Delta shr3$ membranes than using *SHR3* membranes (Fig. 2b; Table 1, top). These results demonstrate that the formation of permease–COPII complexes depends on Shr3. Shr3 was not included in the COPII complex (Fig. 1b), nor is it packaged into COPII vesicles³, so its function is likely to precede permease recognition by the COPII components.

Experiments *in vitro* and *in vivo* demonstrate that vSNAREs, soluble cargo and non-permease integral membrane cargo do not depend on Shr3 for transport^{3,5}. The amount of gp α F and the vSNAREs, Sec22 and Bet1 bound in the Sar1–GST complex from $\Delta shr3$ membranes was similar to the amount bound from *SHR3* membranes (Fig. 2; Table 1, top). A consistent pattern therefore emerges in the specificity and requirements for cargo capture by Sar1–GST and cargo packaging into COPII vesicles.

An independent test of the validity of these complexes was developed using a cleavable crosslinking reagent, dithiobis(succinimidylpropionate) (DSP), followed by denaturing immunoprecipitation. ER membranes incubated with COPII components and treated with DSP were solubilized with SDS and the complexes immunoprecipitated with anti-Sec23 antibody. Recovery of gp α F was dependent on the presence of GTP, Sar1 and DSP, but independent of Shr3 (Fig. 3a; Table 1, bottom). Immunoprecipitation with anti-Sec24 antibodies revealed a similar recovery of cargo-containing complexes (data not shown). Consistent with the Sar1–GST complex results, Gap1–HA was crosslinked in a complex with

COPII in *SHR3* membranes (11.6%) sevenfold more than in $\Delta shr3$ membranes (1.6%). Together with the Sar1–GST data, these experiments show that Sar1, Sec23 and Sec24 are present in a complex with the cargo molecules before vesicle budding.

We then investigated whether coat–cargo complexes persist in transport vesicles. COPII-coated vesicles were treated with DSP and then heated in SDS. Approximately 5% of the gp α F in the vesicles was immunoprecipitated with anti-Sec23 or Sec24 antiserum (Fig. 3b). In control experiments, gp α F was not immunoprecipitated (0.2%) by anti-HA antibody under conditions that precipitated Gap1–HA from the solubilized vesicles (Fig. 3b) or ER membranes (data not shown). Thus the crosslinking conditions were sufficient to couple coat subunits to cargo molecules, but not cargo molecules to one another.

To analyse putative adaptor components, we removed coat proteins from ³⁵S-gp α F-containing COPII vesicles using urea and treated the stripped vesicles with 50 μ M disuccinimidyl glutarate

Table 1 Soluble and membrane cargo in the ER bound to COPII components

Conditions*	Membranes†	DSP (0.5 mM)	gp α F‡ (%)	Gap1–HA‡ (%)
Sar1–GST + Sec23/24	<i>SHR3</i>	–	0.2	0.6
Sar1–GST	<i>SHR3</i>	–	0.01	0.03
Sar1–GST + Sec23/24	$\Delta shr3$	–	0.2	0.1
Sar1 + Sec23/24	<i>SHR3</i>	+	5.4	11.6
Sar1 + Sec23/24	<i>SHR3</i>	–	<0.1	<0.1
Sar1 + Sec23/24 (GDP- β S)	<i>SHR3</i>	+	0.7	1.7
Sec23/24	<i>SHR3</i>	+	0.9	2.6
Sar1	<i>SHR3</i>	+	0.3	0.2
Sar1 + Sec23/24	$\Delta shr3$	+	4.7	1.6

* Incubation conditions for the top section are described in Fig. 2b for the bottom section (gp α F) in Fig. 3a. For the bottom section (Gap1–HA), binding reactions contained 27 μ g perforated spheroplasts from 3-min pulse-labelled cells, 2 μ g Sar1, 5 μ g Sec23/24 and GTP or GDP- β S, as indicated.

† Perforated spheroplasts from PLY129[pPL269] (*SHR3*, Gap1–HA) and PLY134[pPL269] ($\Delta shr3$, Gap1–HA).

‡ In the top section, digitonin-solubilized COPII–cargo complexes were isolated on glutathione–agarose. In the bottom section, SDS-solubilized COPII–cargo complexes were immunoprecipitated with anti-Sec23 antibody. The amount of radiolabelled gp α F and Gap1–HA recovered was compared with the amount of gp α F in total solubilized membranes and Gap1–HA immunoprecipitated from total solubilized membranes, respectively.

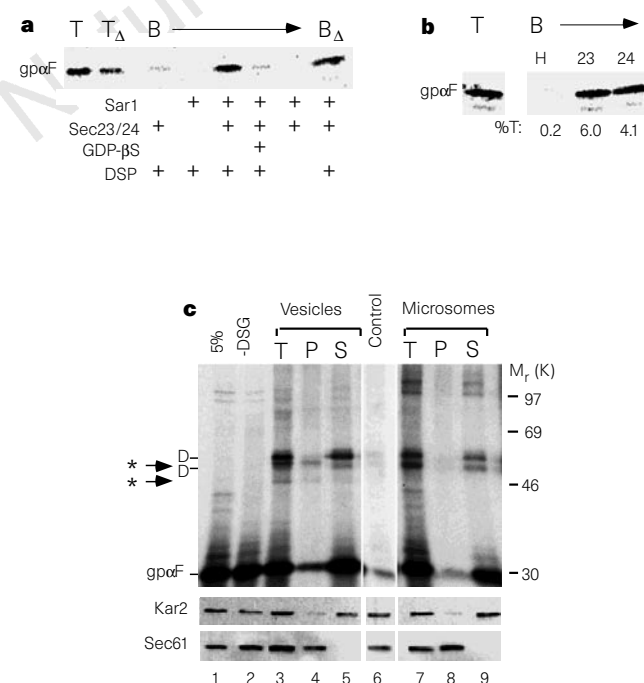


Figure 3 Crosslinking of gp α F in a prebudding complex from ER membranes and ER-derived transport vesicles. **a**, Phosphorimage of 3 μ g total solubilized membranes (T) and bound complexes immunoprecipitated with anti-Sec23 antibody from 300- μ l DSP crosslinked binding reactions (B) containing 17 μ g perforated spheroplasts from ³⁵S-gp α F-loaded PLY129[pPL269] (*SHR3*, Gap1–HA), COPII components, and GTP or GDP- β S as indicated. T_Δ and B_Δ indicate reactions performed with similar amounts of PLY134[pPL269] ($\Delta shr3$, Gap1–HA) membranes. Quantitative analysis is shown in Table 1. **b**, Phosphorimage of ³⁵S-gp α F recovered from total solubilized crosslinked vesicles (T) and bound complexes immunoprecipitated from crosslinking reactions (B). A portion (2%) of the solubilized, crosslinked vesicle preparation was applied directly to the SDS–PAGE (T); the remaining material was divided equally and immunoprecipitated with anti-HA antibody (HA), anti-Sec23 (23), or anti-Sec24 antibody (24). %T: The percentage of gp α F immunoprecipitated compared with the total amount of gp α F in the solubilized vesicles. **c**, ³⁵S-gp α F-loaded microsomes (240 μ g) from RSY445 (lab wild-type strain) were incubated with 1.6 μ g Sar1, 1.6 μ g Sec23/24, and GTP in the presence (lanes 2–5, 7–9) or absence of 20 μ g Sec13/31 (lane 6). The samples were adjusted to 2.5 M urea, crosslinked with DSG (lanes 3–9) and separated by centrifugation. Autoradiograph of samples applied in sample buffer to SDS–PAGE: lane 1, 5% of the microsomes (T) used for vesicle budding reactions (12 μ g); lanes 2, 3 and 6, vesicle-containing fractions; lanes 4 and 8, carbonate-insoluble material (P); lanes 5 and 9, soluble carbonate extract (S); lane 7, 10% (24 μ g) of the crosslinked microsomes (T). The membrane-bound gp α F complexes are indicated by asterisks, the soluble crosslink products representing potential gp α F dimers are labelled D. Bottom, immunoblot control using the integral membrane protein Sec61 and the mainly soluble protein Kar2.

(DSG), a non-cleavable crosslinker. Integral membrane proteins could not be extracted using high pH; soluble proteins were primarily in the supernatant (Fig. 3c, bottom). Two radiolabelled crosslink products of approximate relative molecular mass 46,000 and 56,000 (M_r 48K and 56K) were found to be resistant to carbonate extraction (Fig. 3c), suggesting that they contained gp α F plus transmembrane proteins of ~19K and 27K, respectively. At higher concentrations of crosslinker (500 μ M DSG), these two crosslink products were barely visible, although a radiolabelled crosslink species of 85K was produced (data not shown), possibly representing a dimeric membrane-bound receptor of two ~25K subunits. The three crosslink bands were more abundant in vesicles than in microsomes. The two major crosslink products of approximately 54K and 58K potentially represent dimers of gp α F or underglycosylated gp α F, respectively, and were extractable by high-pH treatment (Fig. 3c). In DSG crosslinking experiments using coated vesicles, the carbonate-inextractable bands were difficult to detect (data not shown), probably owing to their inclusion into aggregates of high M_r . These results indicate that the cross-linked proteins are accessible to proteins on the vesicle surface. Together, these results suggest that at least a portion of the soluble cargo in transport vesicles is complexed to coat proteins through transmembrane adaptors.

We have demonstrated that the COPII components Sar1, Sec23 and Sec24 form a specific complex with integral membrane and soluble vesicle cargo in the ER. The interaction of COPII components with cargo and vesicle proteins, but not with ER chaperones, implies the presence of sorting signals. The absence of a consensus sequence among the transported proteins could mean that different signals have evolved to be decoded by several coat adaptor proteins. We invoke the existence of such an adaptor to link gp α F in the ER lumen to the coat subunits in the cytoplasm. Integral membrane cargo also may require an adaptor to associate with COPII proteins. COPII-soluble cargo complexes were found in isolated vesicles, so the cargo adaptors probably accompany soluble cargo *en route* to the Golgi. It is anticipated that such molecules would be recycled back to the ER. Possible candidates for these functions include redundant members of the p24 family of proteins^{10,11} (one of which we have shown to be part of the COPII prebudding complex) and the ERGIC-53 lectins that cycle between the ER and the Golgi and have been postulated to bind glycosylated cargo during secretion¹³. Isolation of unique recruitment complexes should reveal the identity of the adaptor proteins and other cargo molecules that share a common adaptor. □

Methods

Reagents. Yeast strains and plasmids have been described^{3,5}. pPL327 containing Hip1-Myc⁶ in pRS202 (*URA3*, 2 μ M)¹⁴ was constructed and provided by P. Ljungdahl. Sar1-GST was purified as described² except that the fusion protein was eluted with 10 mM glutathione, 50 mM Tris, pH 7.4, 150 mM NaCl, 0.1% Triton X-100, and dialysed against 20 mM HEPES, pH 6.8, 1 mM magnesium acetate, 150 mM potassium acetate. Polyclonal rabbit antibodies^{4,15-18} and monoclonal mouse antibodies anti-Myc (9E10)¹⁹ and anti-HA (BAbCo) have been described. Standard immunoblotting and immunoprecipitation procedures were performed³.

Sar1-GST assay. Microsomal membranes and perforated spheroplasts were prepared and where noted ³⁵S-prepro- α -factor was translocated *in vitro*³. Membranes were washed with 2.5 M urea in B88 (250 mM sorbitol, 20 mM HEPES, pH 6.8, 5 mM magnesium acetate, 150 mM potassium acetate), then incubated (20 °C, 20 min) with the indicated COPII components (3 μ g GST, 3 μ g Sar1-GST, 5 μ g Sec23/24 and 10 μ g Sec13/31, purified as described^{1,2}) and 0.1 mM guanine nucleotide. An ATP regeneration system¹ (1 $\times$ ATP) was present in some reactions but was not required for binding. The membranes were sedimented (12,000g, 20 min) through 0.5 ml, 0.1 M sucrose/B88-8 (B88, pH 8.0) and the pellet solubilized in 0.5% digitonin/B88-8 (10 min, 25 °C). Insoluble material was removed by centrifugation (12,000g, 10 min), and the supernatant diluted and incubated with glutathione-agarose (GSHA) for

40 min at 4 °C. The beads were washed and bound material was solubilized in 0.5% SDS for 10 min at 55 °C, or in SDS-sample buffer for 3 min at 95 °C (gp α F). Radiolabelled permeases were subsequently immunoprecipitated from the supernatant. Samples were analysed using SDS-PAGE and immunoblotting (ECL, Amersham) or by PhosphorImager (Molecular Dynamics).

Crosslinking. Perforated spheroplasts were washed in urea and incubated (20 min, 20 °C) with COPII components and nucleotide. The membranes were crosslinked (15 min, 20 °C) with 0.5 mM DSP (Pierce) and quenched with 0.4 M ammonium acetate (15 min, 20 °C). Membranes were sedimented through 0.1 M sucrose/B88, solubilized in 1% SDS (10 min, 25 °C), and insoluble material was removed (4,500g, 10 min). The supernatant was immunoprecipitated and incubated with 0.5% SDS and 50 mM DTT for 10 min at 55 °C (Gap1p-HA) or for 3 min at 95 °C in sample buffer with 10% β -mercaptoethanol (β -ME, Bio-Rad) (gp α F). Radiolabelled permeases were immunoprecipitated from the supernatant and heated (10 min, 55 °C) in sample buffer with 10% β -ME. COPII vesicles were isolated from an *in vitro* budding reaction⁶ containing 120 μ g ³⁵S-gp α F-loaded perforated spheroplasts, crosslinked with 0.5 mM DSP, and the reactions quenched as described above. Crosslinked vesicles were centrifuged (70,000g, 15 min) and solubilized in 1% SDS (10 min, 25 °C). Insoluble material was removed (4,500g, 10 min), cross-linked gp α F was immunoprecipitated, and washed beads were heated (3 min, 95 °C) in sample buffer with 10% β -ME. For DSG crosslinking, samples were adjusted to 2.5 M urea, incubated with 50 μ M DSG (Pierce) (20 min, 20 °C), and the reactions quenched with 50 mM glycine (5 min, 20 °C). Microsomes and vesicles were separated by centrifugation (12,000g, 3 min). Microsomes (10 μ g) not containing ³⁵S-gp α F were added to vesicle fractions before the samples were incubated with 0.1 M Na₂CO₃, pH 11.0 (30 min, 0 °C) and separated into soluble and membrane fractions by centrifugation (100,000g, 30 min).

Received 7 July; accepted 9 October 1997.

- Salama, N. R., Yeung, T. & Schekman, R. W. The Sec13p complex and reconstitution of vesicle budding from the ER with purified cytosolic proteins. *EMBO J.* **12**, 4073–4082 (1993).
- Barlow, C. *et al.* COPII: A membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell* **77**, 895–907 (1994).
- Kuehn, M. J., Schekman, R. & Ljungdahl, P. O. Amino acid permeases require COPII components and the ER resident membrane protein Shp3p for packaging into transport vesicles *in vitro*. *J. Cell Biol.* **135**, 585–595 (1996).
- Schimmöller, F. *et al.* The absence of Emp24p, a component of ER-derived COPII-coated vesicles, causes a defect in transport of selected proteins to the Golgi. *EMBO J.* **14**, 1329–1339 (1995).
- Ljungdahl, P. O., Gimeno, C. J., Styles, C. A. & Fink, G. R. *SHR3*: a novel component of the secretory pathway specifically required for localization of amino acid permeases in yeast. *Cell* **71**, 463–478 (1992).
- Barlowe, C. & Schekman, R. SEC12 encodes a guanine-nucleotide-exchange factor essential for transport vesicle budding from the ER. *Nature* **365**, 347–349 (1993).
- Yoshihisa, T., Barlowe, C. & Schekman, R. Requirement for a GTPase-activating protein in vesicle budding from the endoplasmic reticulum. *Science* **259**, 1466–1468 (1993).
- Yeung, T., Barlowe, C. & Schekman, R. Uncoupled packaging of targeting and cargo molecules during transport vesicle budding from the endoplasmic reticulum. *J. Biol. Chem.* **270**, 30567–30570 (1995).
- Kuehn, M. J. & Schekman, R. COPII and secretory cargo capture into transport vesicles. *Curr. Opin. Cell Biol.* **9**, 477–483 (1997).
- Stammes, J. A. *et al.* An integral membrane component of coatamer-coated transport vesicles defines a family of proteins involved in budding. *Proc. Natl Acad. Sci. USA* **92**, 8011–8015 (1995).
- Fiedler, K., Veit, M., Stammes, M. A. & Rothman, J. E. Bimodal interaction of coatamer with the p24 family of putative cargo receptors. *Science* **273**, 1396–1399 (1996).
- Barlow, C., d'Enfert, C. & Schekman, R. Purification and characterization of Sar1p, a small GTP binding protein required for transport vesicle formation from the endoplasmic reticulum. *J. Biol. Chem.* **268**, 873–879 (1993).
- Fiedler, K. & Simons, K. The role of N-glycans in the secretory pathway. *Cell* **81**, 309–312 (1995).
- Sikorski, R. S. & Heiter, P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**, 19–27 (1989).
- Hicke, L., Yoshihisa, T. & Schekman, R. Sec23p and a novel 105 kD protein function as a multimeric complex to promote vesicle budding from the ER. *Mol. Biol. Cell* **3**, 667–676 (1992).
- Stirling, C. J., Rothblatt, J., Hosobuchi, M., Deshaies, R. & Schekman, R. Protein translocation mutants defective in the insertion of integral membrane proteins into the endoplasmic reticulum. *Mol. Biol. Cell* **3**, 129–142 (1992).
- Brodsky, J. L. & Schekman, R. A Sec63p-BIP complex from yeast is required for protein translocation in a reconstituted proteoliposome. *J. Cell Biol.* **123**, 1355–1363 (1993).
- Bednarek, S. Y. *et al.* COPI- and COPII-coated vesicles bud directly from the endoplasmic reticulum in yeast. *Cell* **83**, 1183–1196 (1995).
- Evan, G. I., Lewis, G. K., Ramsay, G. & Bishop, J. M. Isolation of monoclonal antibodies specific for human *c-myc* proto-oncogene product. *Mol. Cell. Biol.* **5**, 3610–3616 (1985).

Acknowledgement. We thank S. Bednarek, M. Bowser and L. Liang for performing experiments; P. Ljungdahl, R. Lesch, M. Pilon and A. Corsi for materials; and J. Campbell, T. Doering and other lab members for suggestions. This work was supported by the Howard Hughes Medical Institute (R.S. and M.K.) and the Cancer Research Fund of the Damon Runyon-Walter Winchell Foundation (M.K.). J.H. was supported by a fellowship from the Deutsche Forschungsgemeinschaft.

Correspondence and requests for materials should be addressed to R.S. (e-mail: schekman@uclink4.berkeley.edu).

Interaction of a G-protein β -subunit with a conserved sequence in Ste20/PAK family protein kinases

Thomas Leeuw*, Cunle Wu*, Joseph D. Schrag†, Malcolm Whiteway*‡, David Y. Thomas*§ & Ekkehard Leberer*||

* Eukaryotic Genetics Group, and † Macromolecular Structure Group, Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Avenue, Montreal, Quebec, Canada H4P 2R2

‡ Department of Biology, §Department of Anatomy and Cell Biology and ||Department of Medicine, McGill University, Montreal, Quebec H3A 2B2, Canada

Serine/threonine protein kinases of the Ste20/PAK family have been implicated in the signalling from heterotrimeric G proteins to mitogen-activated protein (MAP) kinase cascades^{1,2}. In the yeast *Saccharomyces cerevisiae*, Ste20 is involved in transmitting the mating-pheromone signal from the $\beta\gamma$ -subunits (encoded by the *STE4* and *STE18* genes, respectively) of a heterotrimeric G protein to a downstream MAP kinase cascade¹. We have identified a binding site for the G-protein β -subunit ($G\beta$) in the non-catalytic carboxy-terminal regions of Ste20 and its mammalian homologues, the p21-activated protein kinases (PAKs). Association of $G\beta$ with this site in Ste20 was regulated by binding of pheromone to the receptor. Mutations in $G\beta$ and Ste20 that prevented this association blocked activation of the MAP kinase cascade. Considering the high degree of structural and functional

conservation of Ste20/PAK family members and G-protein subunits, our results provide a possible model for a role of these kinases in $G\beta\gamma$ -mediated signal transduction in organisms ranging from yeast to mammals.

The yeast mating-response MAP kinase cascade consists of Ste11 (a MAP or extracellular signal-regulated kinase kinase (MEK) kinase homologue), Ste7 (a MEK homologue) and the partially redundant MAP kinase homologues Fus3 and Kss1 (ref. 1). Activation of this cascade through binding of pheromones to G-protein-coupled receptors induces cellular processes that are typical of differentiating cells, including growth arrest in phase G1 of the cell cycle, differential gene expression, and polarized morphogenesis which leads to the formation of mating-specific projections¹. $G\beta\gamma$ -mediated activation of this cascade involves Ste20 (a MEK kinase kinase) and the MAP kinase scaffolding protein Ste5 (ref. 1). PAKs, a subgroup of mammalian Ste20 homologues, can be activated by either the small G proteins Cdc42 and Rac or by heterotrimeric G proteins in various signalling pathways². The Cdc42 binding domain of Ste20 has been shown to be dispensable for pheromone signalling in yeast, suggesting that activation of Ste20 in response to pheromone occurs in a manner that is independent of Cdc42 (refs 3, 4).

We performed coimmunoprecipitation experiments to analyse the *in vivo* association of Ste20 with an influenza haemagglutinin (HA)-epitope tagged version of $G\beta$ (HA-Ste4). Antibodies to HA-Ste4 precipitated small amounts of Ste20 (Fig. 1a). A roughly five-fold increase in the interaction between Ste4 and Ste20 was observed after just 3 min of pheromone treatment and was maintained for up to 15 min of stimulation (Fig. 1a). The initial induction of Ste20/Ste4 complexes is consistent with the time course described for the stimulation of Far1 (a cyclin inhibitor) and the MAP kinases Fus3 and Kss1 (refs 5, 6) and may be required to activate the MAP kinase cascade for the induction of growth arrest and transcriptional activation. Additional formation of complexes after prolonged

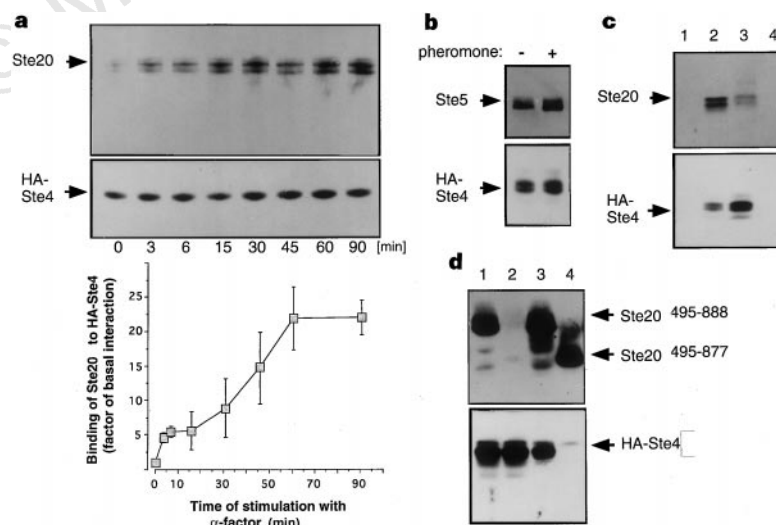


Figure 1 Association of Ste4 with Ste20 and Ste5. **a**, Time course of pheromone-induced Ste20 binding to HA-Ste4. HA-Ste4 expressed from the *STE4* promoter in cells deleted for *STE4* was immunoprecipitated after treatment with 1 μ M α -factor. Relative amounts of Ste20 and HA-Ste4 were determined in western blots (see example in upper and middle panels) and quantified densitometrically (mean values $\pm$ s.d., $n = 3$) (lower panel). **b**, Association of Ste5 with HA-Ste4 expressed from the *STE4* promoter in cells deleted for *STE4*. HA-Ste4 immunoprecipitates from exponentially growing (–) and pheromone-treated (90 min) (+) cells were analysed with antibodies to Ste5 (upper panel) or HA-Ste4 (lower panel). **c**, Overexpression of Ste4 leads to binding to Ste20. HA-Ste4 was overexpressed from the *GAL1* promoter in cells deleted for *STE20* (lane 1) or *STE4* (lanes 2 and 3).

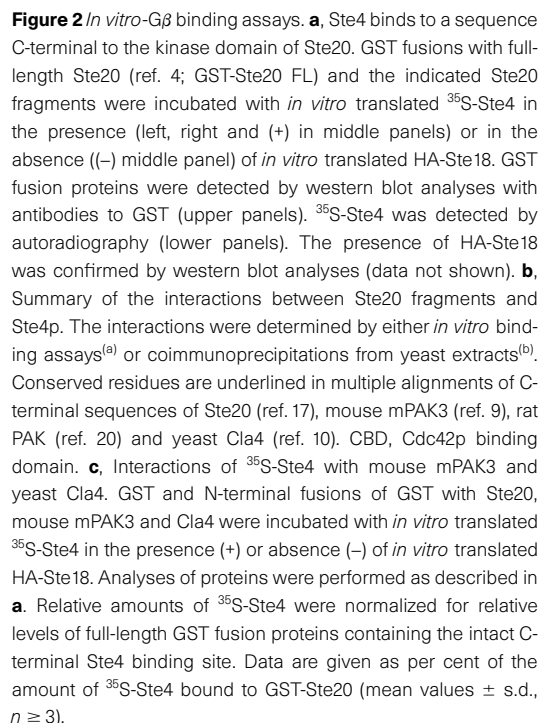
HA-Ste4 expression was suppressed in glucose-containing medium in cells deleted for *STE4* (lane 4). Immunoprecipitates obtained with antibodies to Ste20 (lanes 1 and 2) or the HA-epitope (lanes 3 and 4) were analysed for the presence of Ste20 (upper panel) and HA-Ste4 (lower panel). **d**, Coimmunoprecipitation of HA-Ste4 and Ste20 truncation mutants. HA-Ste4 and Ste20⁴⁹⁵⁻⁸⁸⁸ (lanes 1 and 3) or Ste20⁴⁹⁵⁻⁸⁷⁷ (lanes 2 and 4) truncation mutants were overexpressed from the *GAL1* promoter in cells deleted for *STE20*. HA-Ste4 (lanes 1 and 2) and Ste20 (lanes 3 and 4) immunoprecipitates were analysed for the presence of Ste20 (upper panel) and HA-Ste4 (lower panel) by western blot analyses. Multiple bands of HA-Ste4 and Ste20 represent phosphorylated forms as indicated by phosphatase treatment (data not shown).

When the pheromone response pathway was activated through overexpression of HA-Ste4, Ste5 also formed a complex with Ste4 (ref. 8). This complex was present in cells without an activated pathway when HA-Ste4 was expressed at wild-type levels (Fig. 1b), and the association was not noticeably altered after treatment of cells with pheromone (Fig. 1b), suggesting a constitutive interaction between Ste4 and Ste5. Constitutive activation of the pheromone signalling pathway through overexpression of HA-Ste4 stimulated the association of Ste4 with Ste20 in the absence of pheromone (Fig.

Cells overexpressing the Ste20⁴⁹⁵⁻⁸⁸⁸ fragment were normal in their mating functions, whereas cells overexpressing the Ste20⁴⁹⁵⁻⁸⁷⁷ fragment were defective (Table 1). These results suggest that the region from amino acids 878 to 888 C-terminal to the kinase domain of Ste20 plays an important role in the pheromone response. This region was also required for coimmunoprecipitation with HA-Ste4 (Fig. 1d), underlining the physiological importance of the association between Ste4 and Ste20 in pheromone signalling.

Table 1 Effects of C-terminal truncations on signalling functions of Ste20

Strain YEL206 deleted for *STE20* was transformed with pDH166 (ref. 4), pDH171 and pDH172 carrying either wild-type *STE20* (*STE20*^{WT}) or the *STE20*⁴⁹⁵⁻⁸⁷⁷ and *STE20*⁴⁹⁵⁻⁸⁸⁸ mutant alleles, respectively, under control of the *GAL1* promoter. Mating efficiencies represent mean values $\pm$ s.d. ($n = 3$).



876 to 892 was necessary and sufficient to bind ^{35}S -Ste4 (Fig. 2a). The binding did not depend on the presence of HA-Ste18 (Fig. 2a, c), and HA-Ste18 alone was not able to bind Ste20 (data not shown). As summarized in Fig. 2b, these results, together with data obtained in the immunoprecipitation experiments, suggest that the non-catalytic region from amino acids 878 to 888 of Ste20 represents a binding site for G β .

Consistent with observations that mammalian PAK isoforms can complement the mating defect of yeast cells deleted for *STE20* (ref. 9) and that the G β binding site is conserved in these kinases (Fig. 2b), mouse mPAK3 (ref. 9) bound ^{35}S -Ste4 (Fig. 2c). Ste20 and its closely related isoform Cla4 share a redundant function that is essential for cellular viability in yeast¹⁰. Consistent with observations that only high levels of Cla4 partly complement the mating defect of yeast cells deleted for *STE20* (data not shown) and the Ste4 binding site of Ste20 is not well conserved in Cla4 (Fig. 2b), we found only weak binding of Ste4 to Cla4 (Fig. 2c). These results support the view that residues conserved in the Ste4 binding sites of Ste20 and PAK isoforms (Fig. 2b) contribute to the binding of G β .

Single alterations of the conserved residues S879, S880 or P883 to alanine did not affect the *in vivo* function of Ste20 (data not shown). However, the triple mutant Ste20^{S879A/S880A/P883A}, in which the highly conserved sequence motif SSLxPL was altered to AALxAL, showed strong defects in mating functions (Table 2). Similar defects were also observed for the mutant Ste20^{CLA4}, in which the Ste4 binding site of Ste20 was replaced by the equivalent region of Cla4 (Table 2

and Fig. 2b). No differences were found for the *in vitro* kinase activities of these mutants when compared with wild-type Ste20 (Table 2), and the mutants were found to complement the growth defect of cells deleted for both *STE20* and *CLA4* (data not shown).

However, binding to ^{35}S -Ste4 was strongly reduced when fusions of GST with both mutant versions were analysed in the *in vitro* binding assay (Table 2). Thus, the mating defects of these Ste20 mutants correlated with their reduced ability to bind Ste4.

We identified mutations within two regions of Ste4 which, when overexpressed, inhibited the signalling function of the wild-type protein¹¹. We examined two of these dominant-negative mutations within each region for their effect on the association of Ste4 with either Ste20 or Ste5. The K55E and D62N mutants of Ste4 (ref. 11) were defective in binding to GST-Ste20, whereas binding to GST-Ste5 was normal (Fig. 3a). The inability of these Ste4 mutants to bind Ste20 correlated with their sterile phenotype¹¹. However, the N157H/S175P and Δ F177 mutants which were also found to possess reduced signalling functions¹¹ were able to bind both Ste20 and Ste5, although binding of Ste5 was reduced when compared with binding to wild-type Ste4 (Fig. 3a), suggesting that this region may be involved in the interaction with an as yet unidentified component. Modelling of Ste4 by using the crystal structure of mammalian G β 1 (ref. 12) as a template indicates that the residues predicted to interact with Ste20 are part of an N-terminal α -helix in the region of G β that interacts with G γ (refs 12, 13). However, consistent with the finding that the Ste4^{D62N} mutant interacted normally with Ste18

Table 2 Effect of C-terminal mutations in Ste20 on *in vitro* binding to Ste4, *in vivo* signalling functions and *in vitro* kinase activities

<i>STE20</i> allele	Binding to ^{35}S -Ste4 (%)*	Mating efficiencies (%)†	<i>FUS1::lacZ</i> expression‡	G ₁ arrest†	Shmoo formation†	Kinase activity (%)‡
<i>STE20</i> ^{WT}	100	72.7 ± 6.8	274 ± 41	+	+	100
<i>STE20</i> ^{S879A/S880A/P883A}	15.2 ± 7.8	0.039 ± 0.002	14 ± 1.8	–	–	104 ± 11.2
<i>STE20</i> ^{CLA4}	11.1 ± 9.3	0.016 ± 0.001	11 ± 1.4	–	–	93 ± 10.7
Control	2.1 ± 1.8	<0.005	<0.2 ± 1	–	–	2.5 ± 1.3

Mean values ± s.d. (n = 3).

*GST (control) and fusions of GST with the C-terminal fragments from amino-acids 819 to 939 of wild-type Ste20 (*STE20*^{WT}), the Ste20^{S879A/S880A/P883A} mutant and the Ste20^{CLA4} mutant (in which the sequence encoding amino-acids 879 to 887 of *STE20* was replaced by the sequence of *CLA4* encoding amino-acids 832 to 840 (ref. 10), respectively, were incubated with *in vitro*-translated ^{35}S -Ste4 in the presence of *in vitro* translated HA-Ste18. Data are given as relative levels of bound ^{35}S -Ste4 normalized against binding to wild-type Ste20^{WT} (100%).

†Cells deleted for *STE20* were transformed with pRS313 (control), pSTE20-5 carrying wild-type *STE20* (ref. 17) (*STE20*^{WT}), pBTL150 carrying the Ste20^{S879A/S880A/P883A} mutant, and pBTL151 carrying the Ste20^{CLA4} mutant. Proteins were expressed under control of the *STE20* promoter.

‡*In vitro* kinase activities were determined in immune complexes isolated from YEL206 cells expressing the indicated *STE20* alleles or the inactive *STE20*^{K694R} mutant¹⁴ as a control. Data are given as percentage of MBP phosphorylation by wild-type Ste20.

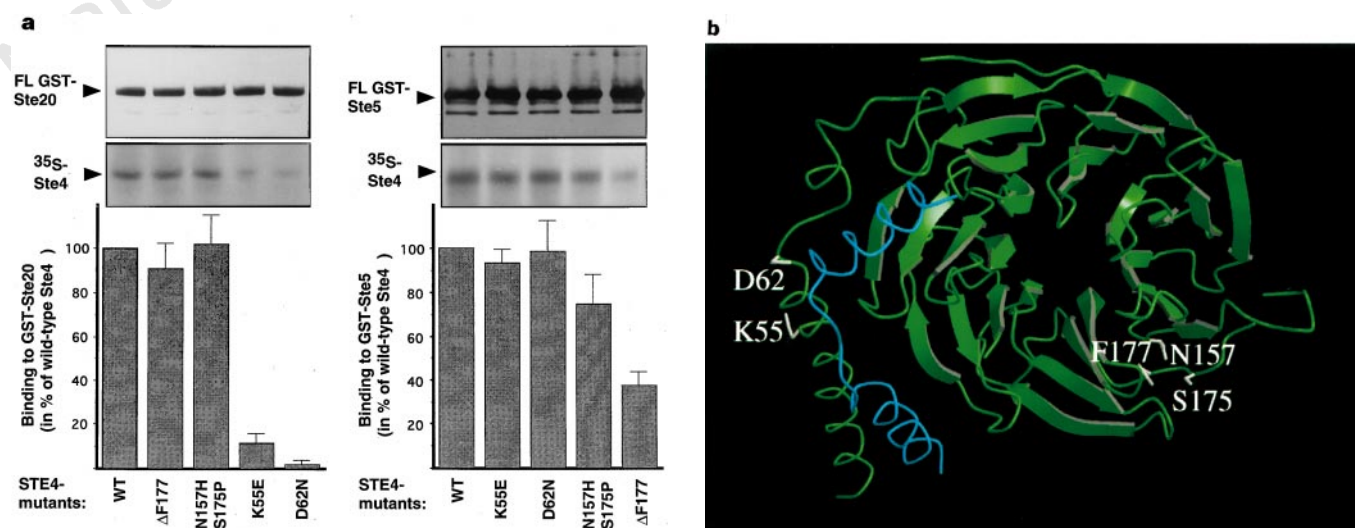


Figure 3 Mutational analysis of the association of Ste4 with Ste20. **a**, Interaction of Ste4 mutants with Ste20 and Ste5. Fusions of GST with Ste20 and Ste5 were incubated in the presence of *in vitro* translated HA-Ste18 with wild-type ^{35}S -Ste4 or the indicated dominant-negative ^{35}S -Ste4 mutants¹¹. Relative amounts of the GST fusion proteins and of ^{35}S -Ste4 were quantified by densitometrical evaluation of

western blots and autoradiographs, respectively. Data are given as percentage of binding of wild-type ^{35}S -Ste4 (mean values ± s.d., n ≥ 3). **b**, Predicted model for the structure of the yeast G β γ dimer. G β (Ste4) is drawn in green and G γ (Ste18) in blue. The white-coloured side chains show residues discovered in a screen for dominant-negative Ste4 mutants¹¹.

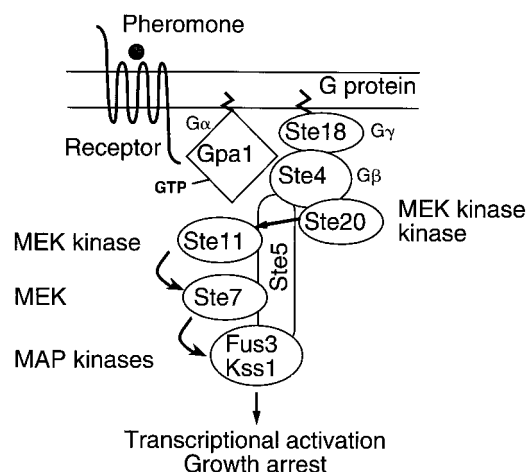


Figure 4 Model for the role of Ste20 in the activation of the pheromone response pathway.

in the two-hybrid system (data not shown), the side chains of these residues are not predicted to be involved in the interaction with $G\gamma$ but rather to be exposed on the cytoplasmic face of the $G\beta\gamma$ structure (Fig. 3b).

Together, our results indicate that transmission of the pheromone signal involves the regulated interaction between the mating-response G-protein β -subunit and a conserved sequence in the Ste20 protein kinase (Fig. 4). Pheromone-induced interaction with Ste4 may bring Ste20 in the vicinity of Ste11 (Fig. 4), which interacts with Ste5 (ref. 1) and can serve as an *in vitro* substrate for Ste20 (ref. 14). Low concentrations of Ste20/ $G\beta$ complexes present in the absence of pheromone may account for the basal signalling levels found in uninduced cells and may guarantee the rapid responsiveness of cells to pheromone^{5,6}. In view of the high degree of conservation of Ste20 family protein kinases², our results suggest that the interaction of these kinases with the β -subunit of heterotrimeric G proteins may contribute to linking Ste20 homologues to G-protein-coupled receptors in other organisms including mammalian cells. □

Methods

Yeast strains and manipulations. *S. cerevisiae* strains used in this study were W303-1A (*MATa ade2 leu2 trp1 ura3 his3 can1*), YEL206 (W303-1A *ste20Δ-3::TRP1*), YEL155 (W303-1A *ste5Δ::TRP1*) and YEL121 (W303-1A *ste4Δ::LEU2*). Mating assays, analysis of mating projection formation, and growth arrest, induction of *FUS1::lacZ* and complementation assays of the growth defect of cells deleted for both *STE20* and *CLA4* were carried out as described^{4,15}.

Construction of plasmids. To construct pBTL110 carrying *HA-STE4* under control of the *STE4* promoter, a fragment from nucleotides -491 to -1 of *STE4* was amplified by polymerase chain reaction (PCR) and cloned into pRS313 (ref. 16). The *Bam*HI fragment of pL55 (ref. 8) was then subcloned downstream of the *STE4* promoter. To create pBTL38 and pBTL65 carrying *STE4* and *HA-STE18* under control of the T3 RNA polymerase promoter, *STE4* and *HA-STE18* were amplified by PCR and ligated into pRS316 and pRS313 (ref. 16), respectively. To create pBTL79, pBTL80, pBTL81 and pBTL82 carrying the *STE4*^{p62N}, *STE4*^{K55E}, *STE4*^{N157H/S175P} and *STE4*^{ΔF177} mutants under control of the T7 RNA polymerase promoter, respectively, the *GAL1* promoter was excised from pGAL-STE4-D62N, pGAL-STE4-K55E, pGAL-STE4-N157H/S175P and pGAL-STE4-ΔF177, respectively¹¹. To create pDH171 and pDH172 carrying the Ste20⁴⁹⁵⁻⁸⁷⁷ and Ste20⁴⁹⁵⁻⁸⁸⁸ fragments under control of the *GAL1* promoter, respectively, these fragments were amplified by PCR and subcloned into pRS313GAL¹⁷.

To create pBTL83, pBTL84, pBTL146 and pBTL147 carrying fusions of GST with the Ste20⁸⁷⁶⁻⁹³⁹, Ste20⁸⁷⁶⁻⁸⁹², Ste20⁸¹⁹⁻⁸⁷⁵ and Ste20⁸¹⁹⁻⁸⁹² fragments,

respectively, fragments were amplified by PCR and subcloned into pGEX-4T-1 (Pharmacia). To create a fusion of GST with full-length Ste5, the *STE5* coding region was amplified by PCR and ligated into pGEX-4T-3 (Pharmacia) to yield pVL50.

Oligodeoxynucleotide-directed mutagenesis of *STE20*. pBTL151 and pBTL150 carrying the *STE20* mutants *STE20*^{CLA4}, in which the sequence encoding amino-acids 879 to 887 of *STE20* was replaced by the sequence of *CLA4* encoding amino-acids 832 to 840 (ref. 10), and *STE20*^{S879A/S880A/P883A}, respectively, under control of the *STE20* promoter, were created by site-directed mutagenesis¹⁸. The mutations were confirmed by sequencing. To create pBTL117 and pBTL118 carrying fusions of GST with the fragments from amino-acids 819 to 939 of the *STE20*^{S879A/S880A/P883A} and *STE20*^{CLA4} mutants, respectively, these fragments were amplified by PCR and subcloned into pGEX-4T-2 (Pharmacia).

Immunochemical procedures. Immunoprecipitation experiments with specific antibodies to the HA-epitope (12CA5 monoclonal and rabbit polyclonal anti-HA antibodies were from Babco, Richmond), Ste20 (ref. 14) and Ste5 (ref. 8) were performed according to standard procedures as described^{8,19}. Antibodies to bacterially produced GST were raised in rabbits and affinity-purified as described¹⁴. For the detection of Ste20 fragments, a secondary sheep antibody specific to rabbit immunoglobulin light chains and a tertiary HRP-conjugated donkey antibody to sheep IgG were obtained from The Binding Site, Lim. Immunoprecipitations were confirmed in at least three independent experiments. For quantification, immunoblots were evaluated by integrating densitometry using an Epson ES 1200-C densitometer and the NIH Image 1.59 software.

In vitro $G\beta$ binding assays. Plasmids were linearized downstream of the termination codons of the respective genes. *In vitro* transcription was performed by using either T3 or T7 RNA polymerase and m⁷G(5')ppp(5')G capped GTP. *In vitro* translation of the resulting mRNA was carried out with ³⁵S-labelled methionine using an *in vitro* translation kit (Promega).

GST fusion proteins were purified on glutathione-Sepharose beads in 20 mM HEPES buffer pH 7.4, containing 100 mM NaCl, 50 mM NaF, 0.5 M sorbitol, 2 mM EDTA, 1 mM Na₃VO₄, 0.1% Triton X-100, 1% BSA (w/v) and a protease inhibitor cocktail, and washed 5 times in phosphate-buffered saline (PBS) (140 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.4) by centrifugation at 1,000g. Proteins (5–10 μg) were incubated with 5 μl of the reticulocyte lysate containing the *in vitro* translated products in 25 μl of PBS for 10 min at 30 °C.

The beads were then washed three times with PBS, separated by SDS-PAGE and analysed by autoradiography of western blots. Results were confirmed in at least three independent experiments. Immunodetection and evaluation of immunoblots and radiographs were then performed as described above. Data obtained for ³⁵S-Ste4 were corrected for relative concentrations of the respective GST fusion proteins.

Miscellaneous. *In vitro* kinase assays were performed as described¹⁴ with myelin basic protein (MBP) as substrate. The structure of yeast Ste4 ($G\beta$) was modelled to the structure of mammalian $G\beta 1$ ¹² using the homology module of Insight (Biosym Inc.). Insertions specific for Ste4 were not considered.

Received 29 July; accepted 13 October 1997.

1. Leberer, E., Thomas, D. Y. & Whiteway, M. Pheromone signalling and polarized morphogenesis in yeast. *Curr. Opin. Genet. Dev.* **7**, 59–66 (1997).
2. Sells, M. A. & Chernoff, J. Emerging from the Pak: the p21-activated protein kinase family. *Trends Cell Biol.* **7**, 162–167 (1997).
3. Peter, M., Neiman, A. M., Park, H.-O., Lohuizen, L. v. & Herskowitz, I. Functional analysis of the interaction between the small GTP binding protein Cdc42 and the Ste20 protein kinase in yeast. *EMBO J.* **15**, 7046–7059 (1996).
4. Leberer, E. *et al.* Functional characterization of the Cdc42p binding domain of yeast Ste20p protein kinase. *EMBO J.* **16**, 83–97 (1997).
5. Chang, F. & Herskowitz, I. Phosphorylation of FAR1 in response to α -factor: a possible requirement for cell-cycle arrest. *Mol. Biol. Cell* **3**, 445–450 (1992).
6. Gartner, A., Nasmyth, K. & Ammerer, G. Signal transduction in *Saccharomyces cerevisiae* requires tyrosine and threonine phosphorylation of FUS3 and KSS1. *Genes Dev.* **6**, 1280–1292 (1992).
7. Wu, C. *et al.* Activation of myosin-I by members of the Ste20p protein kinase family. *J. Biol. Chem.* **271**, 31787–31790 (1996).
8. Whiteway, M. S. *et al.* Association of the yeast pheromone response G protein $\beta\gamma$ subunits with the MAP kinase scaffold Ste5p. *Science* **269**, 1572–1575 (1995).
9. Bagrodia, S., Taylor, S. J., Creasy, C. L., Chernoff, J. & Cerione, R. A. Identification of a mouse p21Cdc42/Rac activated kinase. *J. Biol. Chem.* **270**, 22731–22737 (1995).
10. Cvrckova, F., Devirgilio, C., Manser, E., Pringle, J. R. & Nasmyth, K. Ste20-like protein kinases are required for normal localization of cell growth and for cytokinesis in budding yeast. *Genes Dev.* **9**, 1817–1830 (1995).

11. Leberer, E., Dignard, D., Hougau, L., Thomas, D. Y. & Whiteway, M. Dominant-negative mutants of a yeast G-protein β subunit identify two functional regions involved in pheromone signaling. *EMBO J.* **13**, 4805–4813 (1992).
12. Wall, M. A. *et al.* The structure of the G protein heterotrimer $G_i\alpha 1\beta 1\gamma 2$. *Cell* **83**, 1047–1058 (1995).
13. Sondek, J., Bohm, A., Lambright, D. G., Hamm, H. E. & Sigler, P. B. Crystal structure of a G-protein $\beta\gamma$ dimer at 2.1 Å resolution. *Nature* **379**, 369–374 (1996).
14. Wu, C. L., Whiteway, M., Thomas, D. Y. & Leberer, E. Molecular characterization of Ste20p, a potential mitogen-activated protein or extracellular signal-regulated kinase kinase (MEK) kinase kinase from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **270**, 15984–15992 (1995).
15. Leberer, E. *et al.* Cloning of *Saccharomyces cerevisiae* STE5 as a suppressor of a Ste20 protein kinase mutant: structural and functional similarity of Ste5 to Far1. *Mol. Gen. Genet.* **241**, 241–254 (1993).
16. Sikorski, R. S. & Hieter, P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**, 19–27 (1989).
17. Leberer, E., Dignard, D., Hargus, D., Thomas, D. Y. & Whiteway, M. The protein kinase homologue Ste20p is required to link the yeast pheromone response G-protein $\beta\gamma$ subunits to downstream signalling components. *EMBO J.* **11**, 4815–4824 (1992).
18. Kunkel, T. D., Roberts, J. D. & Zakour, R. A. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Meth. Enzymol.* **154**, 367–382 (1987).
19. Leeuw, T. *et al.* Pheromone response in yeast: association of Bem1p with proteins of the MAP kinase cascade and actin. *Science* **270**, 1210–1213 (1995).
20. Manser, E., Leung, T., Salihuddin, H., Zhao, Z. S. & Lim, L. A brain serine/threonine protein kinase activated by Cdc42 and Rac1. *Nature* **367**, 40–46 (1994).

Acknowledgements. We thank A. Nantel for helpful discussion, D. Hargus, D. Dignard and L. Johnson for excellent technical assistance, R. Tsien for supplying the GFP plasmid and S. Sprang for providing the G β 1 structure coordinates. The NRC publication number for this work is 39976.

Correspondence and request for materials should be addressed to E.L. (e-mail: Ekkehard.Leberer@nrc.ca).

A GTP-exchange factor required for cell orientation

Aljoscha Nern & Robert A. Arkowitz

Division of Cell Biology, MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

The Rho-family of GTPases and their regulators are essential for cytoskeletal reorganization and transcriptional activation in response to extracellular signals^{1,2}. Little is known about what links these molecules to membrane receptors. In the budding yeast *Saccharomyces cerevisiae*, haploid cells respond to mating pheromone through a G-protein-coupled receptor and the $\beta\gamma$ subunit of the G protein³, resulting in arrest of the cell cycle, transcriptional activation, and polarized growth towards a mating partner^{4,5}. The Rho-family GTPase Cdc42 and its exchange factor Cdc24 have been implicated in the mating process^{6,7}, but their specific role is unknown. Here we report the identification of *cdc24* alleles that do not affect vegetative growth but drastically reduce the ability of yeast cells to mate. When exposed to mating pheromone, these mutants arrest growth, activate transcription, and undergo characteristic morphological and actin-cytoskeleton polarization. However, the mutants are unable to orient towards a pheromone gradient, and instead position their mating projection adjacent to their previous bud site. The mutants are specifically defective in the binding of Cdc24 to the G-protein $\beta\gamma$ subunit. Our results demonstrate that the association of an exchange factor and the $\beta\gamma$ subunit of a hetero-trimeric G protein links receptor-mediated activation to oriented cell growth.

Cdc42 and its GDP–GTP exchange factor (GEF) Cdc24 are required for vegetative growth^{8,9} and cell mating^{6,7,10}. The precise function of these proteins in cell mating has been difficult to study because they are essential for viability. We reasoned that if Cdc24 has a specific function in the mating pathway, then *cdc24* alleles should exist that affect cell mating but not vegetative growth. To identify such alleles, a collection of *CDC24* random mutants was screened and three recessive mating mutants, *cdc24-m1*, *-m2* and *-m3*, were isolated. This screen required isolated *cdc24* mutants to be able to support vegetative growth. Further characterization of *cdc24-m* cells revealed that growth between 18 and 37 °C, cell morphology, bud site selection, and actin distribution were similar to wild-type cells (see below and Fig. 1a). The specificity of the *cdc24-m* phenotype is

in contrast to that of all other described *cdc24* mutants, which have strong defects in vegetative growth^{8–10}.

To determine the role of Cdc24 in mating, we examined *cdc24-m1* cells for defects in the mating pathway. The mating efficiency of *cdc24-m1* cells with a wild-type partner was reduced approximately 100-fold compared to wild type (Table 1), and this effect was essentially independent of mating type. When *cdc24-m1* or an enfeebled mater defective in cell fusion were used as mating partners, significantly stronger defects were observed. Such bilateral mating defects suggest impairment in a process such as shmoo (mating projection) formation, orientation or fusion in which a wild-type mating partner can partly compensate for the mutant strain.

Pheromone activation results in several responses, including cell-cycle arrest, MAP-kinase cascade-mediated induction of mating-specific genes, and changes in cell morphology^{4,5}. Pheromone-induced growth arrest determined by halo-assays showed that *cdc24-m1* and wild-type cells responded similarly (Fig. 1b). Furthermore, overexpression of the β -subunit of the yeast heterotrimeric G

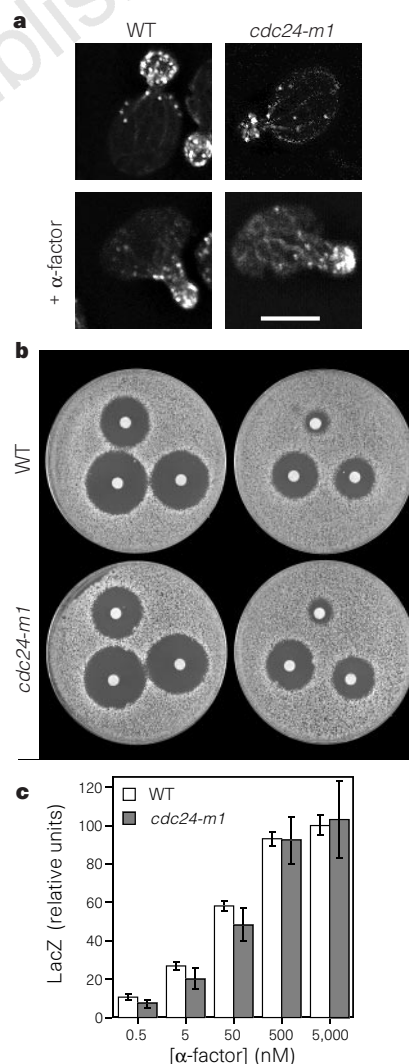


Figure 1 *Cdc24-m1* phenotypes. **a**, Actin cytoskeleton of *cdc24-m1* cells shows polarized distribution. Scale bar 5 μ m. **b**, Pheromone-induced growth arrest is similar in *cdc24-m1* and wild-type (WT) cells. Sterile filter disks spotted with α -factor (1, 0.5, 0.2, 0.1, 0.05 and 0.012 μ g) were placed onto cells in agarose. **c**, MAP-kinase pathway signalling is unaffected in *cdc24-m1*. LacZ activities are the average of 2 experiments (2–3 determinations per experiment) with standard deviation. The WT maximum (29.6 Miller units) was set to 100%.

protein Ste4 from an inducible promoter arrested growth of both *cdc24-m1* and wild-type cells (data not shown). Microscopic examination revealed that identical numbers of wild-type and *cdc24-m1* cells (78%, $n = 1,600$) formed shmoo after 4 h exposure to 10 μ M pheromone. The actin distribution of *cdc24-m1* budding and shmooing cells was also similar to that of wild-type cells (Fig. 1a), demonstrating that the mating defect was not due to an inability to polarize the actin cytoskeleton. The level of pheromone-induced FUS1-lacZ expression, a reporter used to measure the induction of mating-specific genes¹¹, was similar in *cdc24-m1* and wild-type cells (Fig. 1c). However, examination of mating mixtures of *cdc24-m1* and wild-type tester cells showed a greater than tenfold decrease in the number of zygotes, indicating that the *cdc24-m1* defect occurs before cell fusion. Thus *cdc24-m* cells appear normal for cell-cycle

arrest, shmoo formation, actin cytoskeleton polarization, and MAP-kinase signalling, yet are defective at a step before cell fusion.

During mating, polarized growth towards a mating partner requires a pheromone gradient¹², and saturation with pheromone during mating results in random orientation of growth and random selection of a mating partner and hence a decrease in mating efficiency^{13,14}. Wild-type cells showed a 16-fold decrease in mating efficiency in the presence of saturating pheromone (20 μ M), whereas only 10% reduction was observed with *cdc24-m1* cells (Fig. 2a), suggesting that this mutant is unable to orient towards a pheromone gradient during mating. Similar results were observed with *cdc24-m2* and *cdc24-m3* cells. To test directly whether *cdc24-m1* cells are defective in mating projection orientation, their response to an artificial pheromone gradient created by a micropipette was

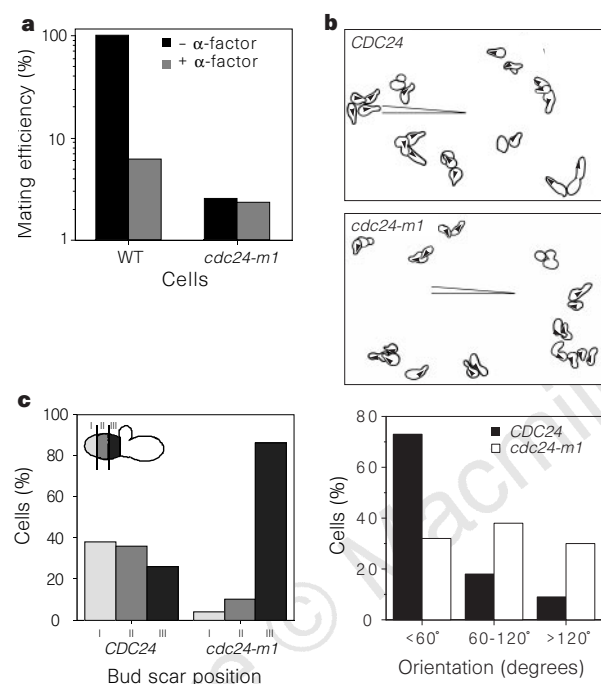


Figure 2 *Cdc24-m1* cells are unable to orient in a pheromone gradient. **a**, Excess pheromone has a negligible effect on *cdc24-m1* mating. *MATa* cells were mated with a wild-type (WT) tester and mating efficiency for *CDC24* (2.8%) was set to 100%. Values are means ($n = 2$). **b**, *Cdc24-m1* cells are unable to orient in a pheromone gradient. A trace of cell shapes after 6–7 h in a pheromone gradient is shown with arrowheads indicating orientation. Quantification of cell projection angle relative to the micropipette (needle) from 4–7 separate experiments ($n = 112$ *CDC24* and 167 *cdc24-m1* cells). The average cosine of the angle of cell projection relative to the micropipette was 0.52 for *CDC24* and -0.02 for *cdc24-m1* cells (a cosine of 1 represents perfect orientation, and 0 random orientation). **c**, *Cdc24-m1* cells position their shmoo adjacent to their bud scar. The position of the bud scar on zygotes was determined for ~120 cells.

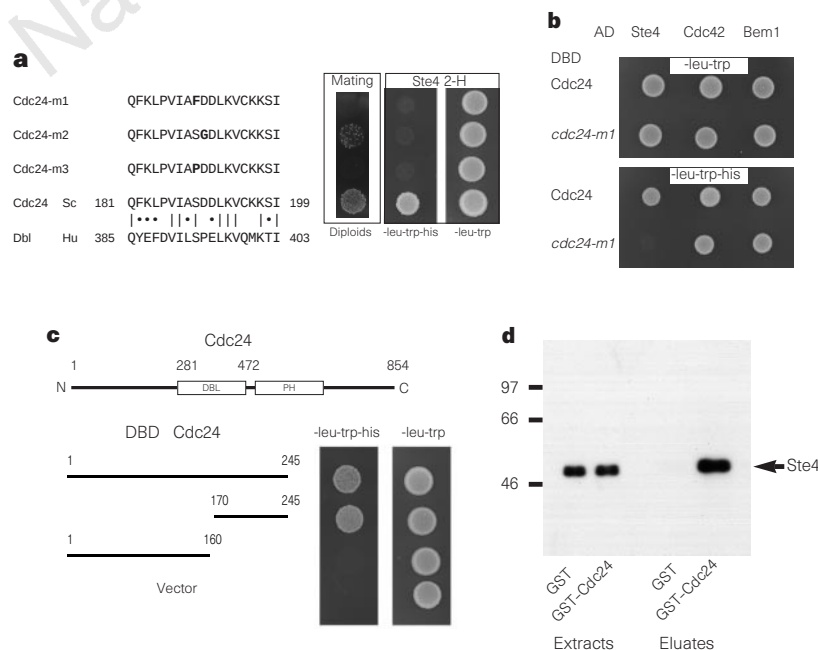


Figure 3 *Cdc24-m* mutants are defective in mating and Ste4 binding. **a**, Location of *Cdc24* mating mutations. Mating patches show diploids from mating with *MATa* wild-type (WT) tester. Ste4 2-H patch growth on -Leu-Trp-His indicates an interaction of *Cdc24* (amino acids 1–288) with Ste4. Similar results were obtained using a LacZ reporter in strain Y187 (relative Miller units 100 for *Cdc24*/Ste4 and 3 for *Cdc24-m1*/Ste4). **b**, Two-hybrid interactions of *Cdc24*. For interactions with Ste4, a fragment of *Cdc24* was used, although full-length *Cdc24* also interacts with Ste4. **c**, Region of *Cdc24* necessary for Ste4 interaction. Numbers refer to *Cdc24* amino acids fused to DBD. **d**, *Cdc24* binds to Ste4 in the absence of other yeast proteins. Mixed bacterial cell extracts (1 eq) containing either His₆Ste4 and GST or GST-*Cdc24* (1–472), and GSH-agarose eluates (800 eq) were separated by SDS-PAGE, immunoblotted and probed with antisera to His₆Ste4. Anti-GST sera showed similar amounts of GST and GST-*Cdc24* in eluates. Owing to proteolysis, His₆Ste4 migrates as a doublet.

examined. The *CDC24* cells oriented growth towards the pheromone source (greater than 70% of cells oriented within 60° of the micropipette), *cdc24-m1* cells did not show a preferred orientation (Fig. 2b). No difference in the sensitivity of wild-type or mutant cells to pheromone was observed.

Although *cdc24-m1* cells oriented randomly in a pheromone gradient, the choice of shmoo site could be dictated by an internal cue, such as the previous bud site^{13,14}. To examine this possibility, the location of the bud scar (in cells with a single bud scar) relative to the neck of the zygote was determined. Wild-type cells showed a random position of their bud scar on the zygotes, whereas 86% of *cdc24-m1* zygotes had a formed from a shmoo adjacent to their previous bud site (Fig. 2c). Taken together, these results establish a specific role for Cdc24 in orientation towards a mating partner.

Sequencing of *cdc24-m* alleles revealed mutations that changed one of two adjacent amino-acid residues (Fig. 3a). Both *cdc24-m1* and *cdc24-m3* cells have a single amino-acid change from Ser 189 to either a Phe or Pro; *cdc24-m2* had two amino-acid substitutions, and subcloning demonstrated that the mutation responsible for the mating defect is Asp to Gly at residue 190. The grouping of these

mutations suggests that this region of Cdc24 is important for an interaction required for oriented growth.

Previous two-hybrid studies have suggested that the amino terminus of Cdc24 might interact with Ste4 (ref. 7), although the *in vivo* significance of this association was unclear. We determined whether Cdc24 mating mutants could interact with Ste4 (Fig. 3a). In contrast to the wild-type Cdc24, the mutants did not show a detectable interaction with Ste4. In agreement with the clustering of the *cdc24-m* mutations, amino-acid residues 170 to 245 of Cdc24 were sufficient for the Ste4 interaction (Fig. 3c), whereas an amino-terminal fragment consisting of the first 160 amino-acid residues, although expressed, failed to interact. Consistent with the interaction between Cdc24 and Ste4 being functionally important, we have isolated mutants in *STE4*, using a two-hybrid screen, which are unable to interact with Cdc24 and are phenotypically similar to *cdc24-m* mutants (A.N. and R.A.A., manuscript in preparation).

To assess the specificity of the defect in the interaction between Ste4 and Cdc24-m1, we investigated the interactions with Cdc42 and Bem1, two proteins known to bind to Cdc24 (refs 15, 16). Bem1 is an SH3-domain protein involved in bud formation and mating¹⁷. Cdc24-m1 was able to interact with both Cdc42 and Bem1 (Fig. 3b), consistent with the absence of an effect of *cdc24-m1* on vegetative growth. Although the *cdc24-m1* phenotype along with the two-hybrid results indicates that the interaction between Cdc24 and Gβ is central to cell orientation, these results do not address whether this interaction is direct or indirect. Gβ typically functions as a complex with the third subunit of a heterotrimeric G protein, Gγ. We therefore determined whether the yeast Gγ, Ste18, was required for the interaction between Cdc24 and Ste4. Deletion of *STE18* abolished the Cdc24 Ste4 two-hybrid interaction (data not shown), suggesting that Cdc24 interacts with the Gβγ complex. To determine whether Cdc24 could bind Ste4 directly, these proteins were expressed in bacteria. Six-histidine-tagged Ste4 specifically bound to a Cdc24 fusion protein with glutathione S-transferase (GST) (Fig. 3d). These results demonstrate that Cdc24 can bind Gβ directly in the absence of any other yeast proteins. We attribute the requirement for Gγ in the two-hybrid assays to its stabilization of Gβ¹⁸.

Pheromone-receptor activation results in dissociation of Gβγ from Gα at the receptor. Our results indicate that the orientation defect in *cdc24-m* cells is due to a specific defect in the Cdc24 Gβγ interaction. This suggests a model in which binding of Cdc24 to Gβγ results in recruitment (to the vicinity of the receptor) or activation of Cdc24, and that this local concentration of activated Cdc24 is required for oriented growth towards a pheromone gradient (Fig. 4). In the absence of this recruitment or activation, a site adjacent to the previous bud site appears to function as a default site for shmoo formation. Our results, together with previous studies implicating Cdc24 in bud site selection⁸, suggest that Cdc24 acts as a crucial component required for the selection of both bud and shmoo sites, perhaps functioning as a kind of molecular selector switch between internal signals for bud site selection and external signals for shmoo site selection. It is likely that local activation of Cdc24 recruits and activates the Rho GTPase Cdc42, which could then interact with downstream targets required for orientation of the cytoskeleton. Cdc42 interactions with the protein kinase Ste20 (refs 19, 20) are not necessary for cell orientation²⁰, suggesting that other targets of Cdc42 are required for oriented growth towards a mating partner.

Cdc24 belongs to a diverse family of GEFs that includes many mammalian proto-oncogenes². This group of proteins shares a conserved region consisting of a Dbl domain (named after the human proto-oncogene *DBL*) followed by a pleckstrin-homology domain (PH). Sequence comparison revealed similarity between a small stretch of amino acids flanking the *cdc24* mating mutations and Dbl (Fig. 3a). Our results indicate that an association between Cdc24 and Gβγ links pheromone receptor activation to shmoo

Table 1 *Cdc24-m1* is defective in cell mating

Strain	Tester	Mating efficiency (%)
<i>CDC24 MATα</i>	<i>MATa</i> wild type	100 (21)
<i>cdc24-m1 MATα</i>	<i>MATa</i> wild type	0.5 (0.2)
<i>CDC24 MATa</i>	<i>MATα</i> wild type	100 (20)
<i>cdc24-m1 MATa</i>	<i>MATα</i> wild type	3.8 (1.6)
<i>CDC24 MATa</i>	<i>MATα Δfus1 Δfus2</i>	100 (17)
<i>cdc24-m1 MATa</i>	<i>MATα Δfus1 Δfus2</i>	≤0.02
<i>CDC24 MATa</i>	<i>MATα CDC24</i>	100 (18)
<i>cdc24-m1 MATa</i>	<i>MATα cdc24-m1</i>	≤0.0006

Mating efficiencies are the number of diploid cells divided by the total cells with *CDC24* wild type set to 100%. The values are means of 4 determinations with standard deviation given in brackets. Absolute mating efficiency was 14–15% with *MATa* and *MATα* testers, 1.8% with *Δfus1 Δfus2* tester, and 3.4% with *CDC24* tester.

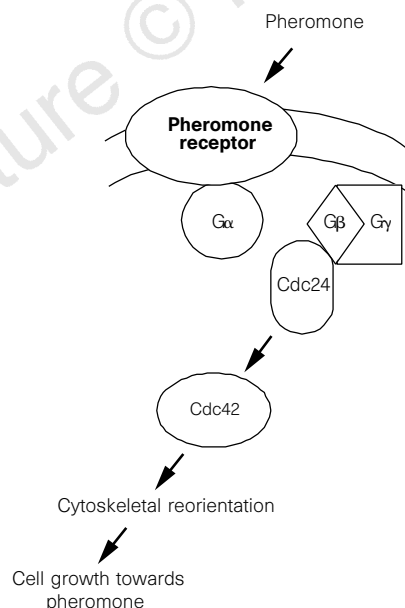


Figure 4 Model of the signal transduction pathway required for cell orientation. For clarity we have omitted components of the MAP-kinase cascade. The role of Cdc42 (a rho-family GTPase) in cell orientation is speculative. Pheromone binds the pheromone receptor (Ste2 or Ste3) resulting in the dissociation of Gα (Gpa1) from Gβγ (Ste4/Ste18). Direct binding of Cdc24 to Gβγ (in the vicinity of the receptor) activates or recruits Cdc24, which is necessary for oriented growth towards a mating partner.

orientation. We propose that other GEFs, such as the proto-oncogene *DBL*, provide a similar connection between G-protein-coupled receptor activation and polarized cell growth. □

Methods

General techniques. Strains were constructed using standard techniques²¹. All constructs were verified by DNA dye terminator cycle sequencing (ABI 377 sequencer).

Strains. pRS414CDC24 contains the *CDC24* ORF including 258 base-pairs (bp) upstream of ATG. Oligonucleotide-directed mutagenesis was used to introduce silent base changes that resulted in the following ten new restriction sites in *CDC24*: *NheI* (bp -12), *KasI* (bp 283), *AatII* (bp 681), *PstI* (bp 1207), *RsrII* (bp 1369), *BstEII* (bp 1426), *XhoI* (bp 1758), *MluI* (bp 1963), *SalI* (bp 2061), *BamHI* (bp 2485). RAY410 (*MAT α* , *leu2*, *cdc24::LEU2*, *ade2*, *lys2*, *his3*, *trp1*, *ura3*, pEG(KT)CD24) was derived from the diploid YOC380 (ref. 22) which was transformed with pEG(KT)CDC24 (ref. 23) and sporulated. RAY950 is isogenic to RAY410 but has pRS416GalHis₆CDC24 as a rescuing plasmid. RAY928 (*MAT α* , *leu2-3*, *112*, *ura3-52*, *his3- Δ 200*, *trp1- Δ 901*, *lys2-801*, *suc2- Δ 9*, *cdc24 Δ -1::HIS5* pEG(KT)CDC24) and RAY931 (same as RAY928 but *MAT α* , *ade2*, *LYS2*) were made by transformation of SEY6210 and 6211 with pEG(KT)CDC24 followed by polymerase chain reaction (PCR)-based gene disruption of *CDC24*. The *CDC24* ORF was replaced with *Schizosaccharomyces pombe HIS5* (ref. 24), flanked by *LoxP* sites. Replacement of *CDC24* in SEY6211 with a PCR-generated integration cassette consisting of *TRP1* fused to 343 bp of *CD24* promoter followed by 1,704 bp of *CD24* or *cdc24-m1* ORF was used to construct RAY1034 or RAY1035, respectively.

Isolation of mating mutants. Base pairs -105 to 789 of *CD24* were mutagenized by PCR²⁵ using pRS414CDC24 as a template and cloned in *NheI*, *AatII*-cut pRS414CDC24. This library was transformed into RAY950 and selected on glucose plates which repress *CDC24* expression. Colonies were then replica-plated onto a tester strain lawn, incubated at 30°C for 3 h, and subsequently replica-plated to select for diploids. Plasmids were isolated from mating mutants, transformed into RAY950, RAY928 and RAY931, and rescuing plasmids removed by 5-FOA counter-selection.

Phenotypic analyses. Quantitative matings¹⁰, matings in the presence of saturating pheromone¹³, halo-assays²⁶ using *ss11::URA3* strains, and *Fus1lacZ* measurements with pSG231 (ref. 11) were performed as described. Halo-assays showed that *MAT α* and *MAT α cdc24-m1* cells secreted α -factor and α -factor, respectively. Actin was visualized with rhodamine phalloidin²⁷ on a Biorad-MRC-600 confocal microscope, and pictures are projections of 4–6 0.5- μ m z-series steps. For α -factor treatment, cells were incubated with 5 μ M α -factor for 2 h. RAY1034 and RAY1035 cells were used to determine bud scar positions on zygotes¹⁴ visualized with Calcofluor²⁸. Similar results were observed with the position of the bud scar on shmoo. Direct measurement of cell orientation in a pheromone gradient was carried out essentially as described¹². A pheromone gradient was generated using a micropipette filled with 80 μ M α -factor injected at 105 kPa into 1 ml of YEPD media layered on top of cells embedded in 2% LMP agarose. Cell shape was recorded by video microscopy on a heated stage at 35°C for 4–7 h and data analysis was from traced cell outlines¹⁴. Mating projections were formed at the same pheromone concentrations and budding, that is, non-responding cells were seen at similar distances from the micropipette in both strains.

Two-hybrid methods. *STE4*, *BEM1* (372–551 amino acids), *CDC42*[C178S] and *CDC24/cdc24-m1* (1–288, 1–160 and 170–245 amino acids) were cloned by PCR into pGAD424 (AD, *GAL4* activation domain) or pAS1 (DBD, *GAL4* DNA binding domain). Plasmids were transformed into HF7c. For determination of *STE18* requirement, PCR-based gene disruption was carried out in PJ69-4A (*MAT α* , *trp1-901*, *leu2-3*, *112*, *ura3-52*, *his3-200*, *gal4 Δ* , *gal80 Δ* , *GAL2-ADE2*, *LYS2::GAL1-HIS3*, *met2::GAL7-LacZ*)²⁹, replacing the entire *STE18* ORF with *Kluyveromyces lactis URA3* (ref. 30). For all two-hybrid experiments, equal amounts of transformants were spotted on SC-Leu-Trp and SC-Leu-Trp-His plates, identical results were obtained with at least four transformants, and for Δ *ste18* two independent deletion strains. All strains for two-hybrid analyses expressed similar amounts of AD and DBD fusion proteins of the expected sizes, as determined by SDS-PAGE and immunoblotting. None of the DBD fusions showed any self-activation using two different non-interacting AD fusions.

In vitro binding studies. A fragment of *CDC24* (amino acids 1–472) in pGEX-2T (Pharmacia) and *STE4* in pTrcA (Invitrogen) were expressed in *E. coli*. Cells were resuspended in buffer A (PBS, 0.1% TX-100, PMSF, leupeptin, chymostatin, pepstatin, aprotinin) and lysed by snap freezing in liquid nitrogen followed by sonication. Insoluble material was removed by centrifugation (10,000g). Mixed supernatants (denoted cell extracts) containing His₆Ste4 and GSTCdc24 fusions were incubated with GSH-agarose (Sigma) at 4°C for 1 h. Resin was washed 3 times with buffer A. Resin samples (referred to as eluates) and extracts were analysed by SDS-PAGE, immunoblotting probed with Omni-probe antisera (Santa Cruz), and visualized with enhanced chemiluminescence (Amersham). GST-Cdc24 (amino acids 1–127), similar to GST, did not bind His₆Ste4. Similar results were observed in 5 independent experiments.

Received 4 July; accepted 4 November 1997.

- Machesky, L. M. & Hall, A. Rho: a connection between membrane receptor signaling and the cytoskeleton. *Trends Cell Biol.* **6**, 304–310 (1996).
- Cerione, R. A. & Zheng, Y. The Dbl family of oncogenes. *Curr. Opin. Cell Biol.* **8**, 216–222 (1996).
- Whiteway, M. *et al.* The *STE4* and *STE18* genes of yeast encode potential β -subunits and γ -subunits of the mating factor receptor-coupled G protein. *Cell* **56**, 467–477 (1989).
- Sprague, G. F. Jr & Thorner, J. W. in *The Molecular and Cellular Biology of the Yeast Saccharomyces* (eds Jones, E. W., Pringle, J. R. & Broach, J. R.) 657–744 (Cold Spring Harbor Laboratory Press, NY, 1992).
- Leberer, E., Thomas, D. Y. & Whiteway, M. Pheromone signalling and polarized morphogenesis in yeast. *Curr. Opin. Genet. Dev.* **7**, 59–66 (1997).
- Simon, M.-N. *et al.* Role for the Rho-family GTPase Cdc42 in yeast mating pheromone signal pathway. *Nature* **376**, 702–705 (1995).
- Zhao, Z. S., Leung, T., Manser, E. & Lim, L. Pheromone signalling in *Saccharomyces cerevisiae* requires the small GTP-binding protein Cdc42p and its activator *CDC24*. *Mol. Cell Biol.* **15**, 5246–5257 (1995).
- Sloat, B. F., Adams, A. & Pringle, J. R. Roles of the *CDC24* gene product in cellular morphogenesis during the *Saccharomyces cerevisiae* cell cycle. *J. Cell Biol.* **89**, 395–405 (1981).
- Adams, A. E., Johnson, D. I., Longnecker, R. M., Sloat, B. F. & Pringle, J. R. *CDC42* and *CDC43*, two additional genes involved in budding and the establishment of cell polarity in the yeast *Saccharomyces cerevisiae*. *J. Cell Biol.* **111**, 131–142 (1990).
- Chenevert, J., Valtz, N. & Herskowitz, I. Identification of genes required for normal pheromone-induced cell polarization in *Saccharomyces cerevisiae*. *Genetics* **136**, 1287–1296 (1994).
- Trueheart, J., Boeke, J. D. & Fink, G. R. Two genes required for cell fusion during yeast conjugation: evidence for a pheromone-induced surface protein. *Mol. Cell Biol.* **7**, 2316–2328 (1987).
- Segall, J. E. Polarization of yeast cells in spatial gradients of α mating factor. *Proc. Natl Acad. Sci. USA* **90**, 8332–8336 (1993).
- Dorer, R., Pryciak, P. M. & Hartwell, L. H. *Saccharomyces cerevisiae* cells execute a default pathway to select a mate in the absence of pheromone gradients. *J. Cell Biol.* **131**, 845–861 (1995).
- Valtz, N., Peter, M. & Herskowitz, I. *FAR1* is required for oriented polarization of yeast cells in response to mating pheromones. *J. Cell Biol.* **131**, 863–873 (1995).
- Zheng, Y., Cerione, R. & Bender, A. Control of the yeast bud-site assembly GTPase Cdc42. Catalysis of guanine nucleotide exchange by Cdc24 and stimulation of GTPase activity by Bem3. *J. Biol. Chem.* **269**, 2369–2372 (1994).
- Zheng, Y., Bender, A. & Cerione, R. A. Interactions among proteins involved in bud-site selection and bud-site assembly in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **270**, 626–630 (1995).
- Chenevert, J., Corrado, K., Bender, A., Pringle, J. & Herskowitz, I. A yeast gene (*BEM1*) necessary for cell polarization whose product contains two SH3 domains. *Nature* **356**, 77–79 (1992).
- Hirschman, J. E., DeZutter, G. S., Simonds, W. F. & Jenness, D. D. The G $\beta\gamma$ complex of the yeast pheromone response pathway—subcellular fractionation and protein–protein interactions. *J. Biol. Chem.* **272**, 240–248 (1997).
- Peter, M., Neiman, A. M., Park, H. O., Van Lohuizen, M. & Herskowitz, I. Functional analysis of the interaction between the small GTP binding protein Cdc42 and the Ste20 protein kinase in yeast. *EMBO J.* **15**, 7046–7059 (1996).
- Leberer, E. *et al.* Functional characterization of the Cdc42p binding domain of yeast Ste20p protein kinase. *EMBO J.* **16**, 83–97 (1997).
- Rose, M. D., Winston, F. & Hieter, P. *Methods in Yeast Genetics: A Laboratory Course Manual* (Cold Spring Harbor Laboratory Press, NY, 1991).
- Miyamoto, S. *et al.* A DBL-homologous region of the yeast *CLS4/CDC24* gene product is important for Ca(2+)-modulated bud assembly. *Biochem. Biophys. Res. Commun.* **181**, 604–610 (1991).
- Mitchell, D. A., Marshall, T. K. & Deschenes, R. J. Vectors for the inducible overexpression of glutathione-S-transferase fusion proteins in yeast. *Yeast* **9**, 715–722 (1993).
- Wach, A., Brachat, A., Alberti-Segui, C., Rebischung, C. & Philippsen, P. Heterologous *HIS3* marker and GFP reporter modules for PCR-targeting in *Saccharomyces cerevisiae*. *Yeast* **13**, 1065–1075 (1997).
- Fromant, M., Blanquet, S. & Plateau, P. Direct random mutagenesis of gene-sized DNA fragments using polymerase chain-reaction. *Anal. Biochem.* **224**, 347–353 (1995).
- Sprague, G. F. Assay of yeast mating reaction. *Methods Enzymol.* **194**, 77–93 (1991).
- Adams, A. E. & Pringle, J. R. Staining of actin with fluorochrome-conjugated phalloidin. *Methods Enzymol.* **194**, 729–731 (1991).
- Pringle, J. R. Staining of bud scars and other cell wall chitin with calcofluor. *Methods Enzymol.* **194**, 732–733 (1991).
- James, P., Halladay, J. & Craig, E. A. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**, 1425–1436 (1996).
- Langle-Ronault, F. & Jacobs, E. A method for performing precise alterations in the yeast genome using a recyclable selectable marker. *Nucleic Acids Res.* **23**, 3079–3081 (1995).

Acknowledgements. We thank Y. Anraku, R. Deschenes, G. Fink, E. Jacobs, P. James and P. Philippsen for strains and plasmids; M. Bassilana, M. Bienz, T. Bretscher, H. Pelham and S. Munro for comments and suggestions; and N. Lowe for assistance with sequencing. This work was supported by the Medical Research Council. A.N. was supported by fellowships from the Fonds der Chemischen Industrie and the EC.

Correspondence and requests for materials should be addressed to R.A.A. (e-mail: ra2@mrc-lmb.cam.ac.uk).

Structure of the $\alpha\beta$ tubulin dimer by electron crystallography

Eva Nogales, Sharon G. Wolf* & Kenneth H. Downing

Life Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

The $\alpha\beta$ tubulin heterodimer is the structural subunit of microtubules, which are cytoskeletal elements that are essential for intracellular transport and cell division in all eukaryotes. Each tubulin monomer binds a guanine nucleotide, which is non-exchangeable when it is bound in the α subunit, or N site, and exchangeable when bound in the β subunit, or E site. The α - and β -tubulins share 40% amino-acid sequence identity, both exist in several isotype forms, and both undergo a variety of post-translational modifications¹. Limited sequence homology has been found with the proteins FtsZ² and Misato³, which are involved in cell division in bacteria and *Drosophila*, respectively. Here we present an atomic model of the $\alpha\beta$ tubulin dimer fitted to a 3.7-Å density map obtained by electron crystallography of zinc-induced tubulin sheets. The structures of α - and β -tubulin are basically identical: each monomer is formed by a core of two β -sheets surrounded by α -helices. The monomer structure is very compact, but can be divided into three functional domains: the amino-terminal domain containing the nucleotide-binding region, an intermediate domain containing the Taxol-binding site, and the carboxy-terminal domain, which probably constitutes the binding surface for motor proteins.

In the presence of zinc ions, purified tubulin assembles into two-dimensional sheets that are ideal samples for electron crystallography studies⁴. In these sheets, protofilaments appear to be similar to those in microtubules, but associated in an antiparallel fashion. The zinc-induced tubulin sheets used here are cold-labile and require GTP for assembly. Addition of taxol stabilizes the sheets against low-temperature depolymerization and ageing⁴, an effect similar to that on microtubules. Taxol binds to a single site on the dimer in the sheets, near lateral contacts between protofilaments⁵.

We have previously described three-dimensional density maps of tubulin at 6.5 and 4 Å (refs 5, 6). The present model has been built into a 3.7-Å map derived from a data set that includes 93 electron-diffraction patterns (providing structure factor amplitudes) and 159 images (providing experimental phases). Table 1 and Fig. 1 summarize the data. The high quality of the phases produced a clean map, with well defined connectivity, which is readily interpretable in terms of secondary-structure elements. The present model of tubulin has been built in the raw density map without refinement and includes all but the last 10 and 18 C-terminal residues of α - and β -tubulin, respectively. Figure 2 shows the density map and the model in several regions of the dimer.

The density maps for α - and β -tubulin are almost superimposable. Differences are limited to the length and conformation of some loops, very slight displacements (~1 Å) of some of the secondary-structure elements, and differences in side-chain densities. Owing to the similarity between monomers, our description and comments on the model apply to both α - and β -tubulin unless indicated otherwise. Residue numbers correspond to the aligned sequences of α - and β -tubulin as shown in Fig. 3, and include gaps in the sequence of β -tubulin.

Figure 4 shows the ribbon diagram of the tubulin dimer model (see later for 'dimer' definition). The core of the structure contains

two β -sheets of 6 and 4 strands, flanked by 12 α -helices. Although there is no clear division of the density into domains in the map, it makes sense functionally to divide the structure of each monomer into three sequential domains.

The N-terminal domain includes residues 1–205 and forms a Rossmann fold, which is typical of nucleotide-binding proteins, in which parallel β -strands alternate with α -helices. Helices H1 and H2 are on one side of the sheet, whereas helices H3, H4 and H5 are on the other. Strands B2 and B3 are poorly defined in the unrefined density map. The residues in the loops connecting H1 and B2, and H2 and B3, are included for completeness, but were built in very weak density. These segments are long predicted loops on the putative inside surface of the microtubule, and correspond to a region of the sequence that can accommodate insertions and deletions⁷.

Strand B6 leads to the intermediate domain, residues 206–381, containing a mixed β -sheet and five surrounding helices. The domain starts with helices H6 and H7, followed by a long loop and helix H8 at the longitudinal interface between monomers. B7 is a long β -strand that interacts with the β -sheet in the N-terminal domain. The loop connecting B7 and H9 is more ordered in α -tubulin, where it is involved in strong lateral contacts. Following a long loop

Table 1 Electron crystallographic data

Two-dimensional crystals	
Two-sided plane group	P12 ₁
Unit cell	$a = 80, b = 92 \text{ Å}$
Sampling thickness	$c = 90 \text{ Å}$
Sample thickness	62 Å
Experimental data set	
Resolution cut-off	3.7 Å
Number of structure factors	12,000
Electron diffraction	
Number of patterns by tilt angle	18 (0°), 57 (45°), 19 (55°)
R_{sym}	19%
R_{merge}	25%
Image/phase data	
Number of images by tilt angle	12 (0°), 51 (45°), 86 (60°)
Phase residual by resolution zones	36° (5–4 Å), 46° (4–3.7 Å)

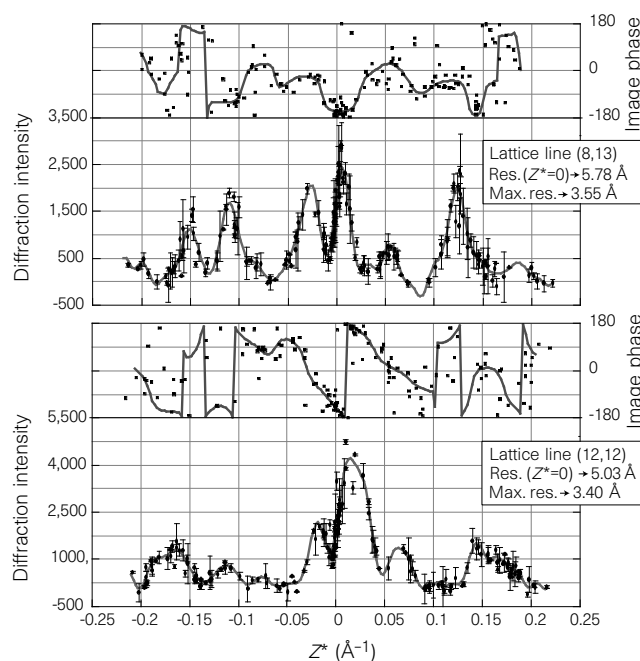


Figure 1 Experimental phase and intensity data and fitted curves for two representative reciprocal lattice lines. Error bars for the intensities are Friedel-pair differences. The resolution of each lattice line at the equator ($z^* = 0$) and its furthest point are indicated.

* Present address: Electron Microscopy Unit, Weizmann Institute of Science, Rehovot 76100, Israel.

come B8 and H10, which unravels at its C terminus in α -tubulin, followed by a very well defined short loop into B9. Finally, the loop between B9 and B10 includes an 8-residue insertion in the α -subunit which occludes the site that in β is occupied by taxol. The present model is compatible with the observation that C241 and C356 in the β -subunit can be crosslinked⁸ (~8 Å apart in the model; Fig. 4b).

The C-terminal domain is formed by helices H11 and H12. These helices overlay the previous domains, sitting on the surface of the molecule that we have identified as the outside surface in the microtubule⁹ (Fig. 4, legend). They are probably involved in the binding of MAPs and motor proteins. The loop connecting H11 and H12 is important for the interaction with the next monomer along the protofilament.

The last C-terminal residues of each monomer, missing from the model, correspond to the hypervariable part of the sequence where most of the differences between isotypes and across species occur. As the tubulin preparation we used included several α and β isotypes that are found in bovine brain, the inhomogeneity of the sample could contribute to the poor visibility in this part of the map. However, we do not see any significant difference in projection maps among α III, α II, and undifferentiated tubulin (our unpublished results). This tubulin segment is highly acidic in both monomers and likely to be disordered, irrespective of the isotype composition.

The nucleotide in tubulin is positioned at the base of the Rossman fold. The B1–H1, B2–H2 and B3–H3 loops, and the glycine-rich B4–H4 loop, appear to contact the phosphates. The

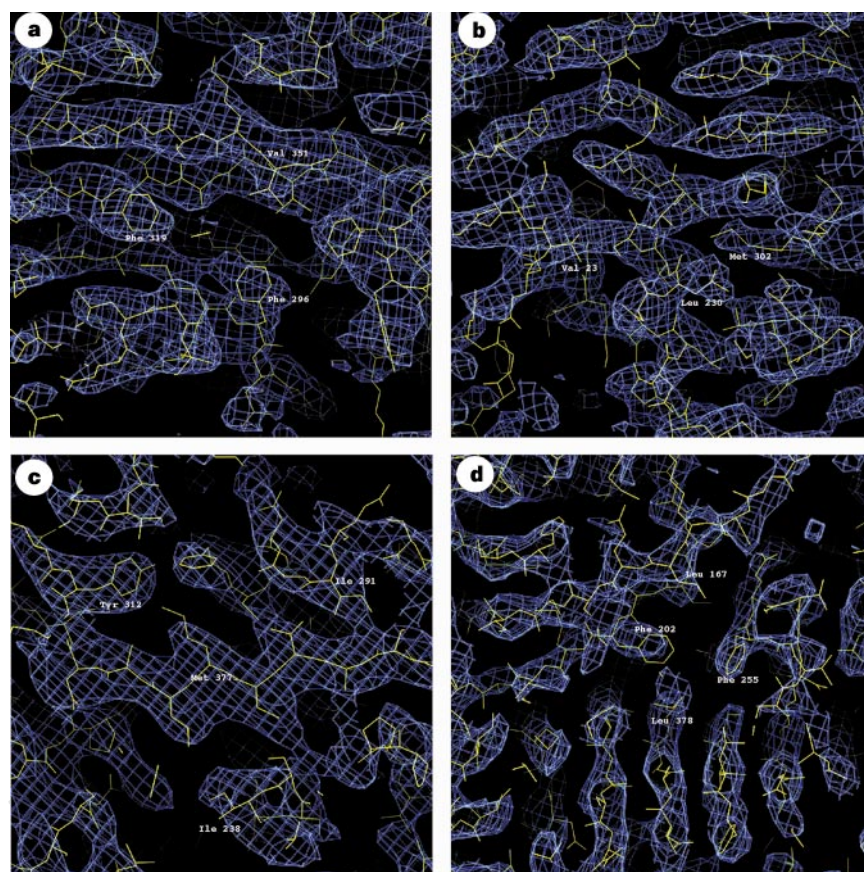


Figure 2 Sections of the experimental density map with the fitted model for different regions in the α - and β -tubulin molecules. **a, b**, Sections from β -tubulin; **c, d**, sections from α -tubulin. The map was calculated with a 3.7-Å cutoff. In the direction perpendicular to the crystals, the resolution is less than 3.7 Å, owing to the limit of 60° in tilt angle, but still sufficient that individual β -strands can be distinguished. The initial assignment of the sequence to the backbone trace was based mostly on the positions of aromatic side chains and on the connectivity of the density. Comparing corresponding densities in the α - and β -subunits, and relating the differences and similarities to those in their amino-acid sequences, were extremely helpful in tracing the chain. The sequence used for the model, in the absence of that from bovine brain tubulin, corresponds to porcine brain tubulin²⁸. The numbering of residues is based on the alignment of the α - and β -tubulin sequences and so includes gaps in the sequence of β -tubulin (Fig. 3).

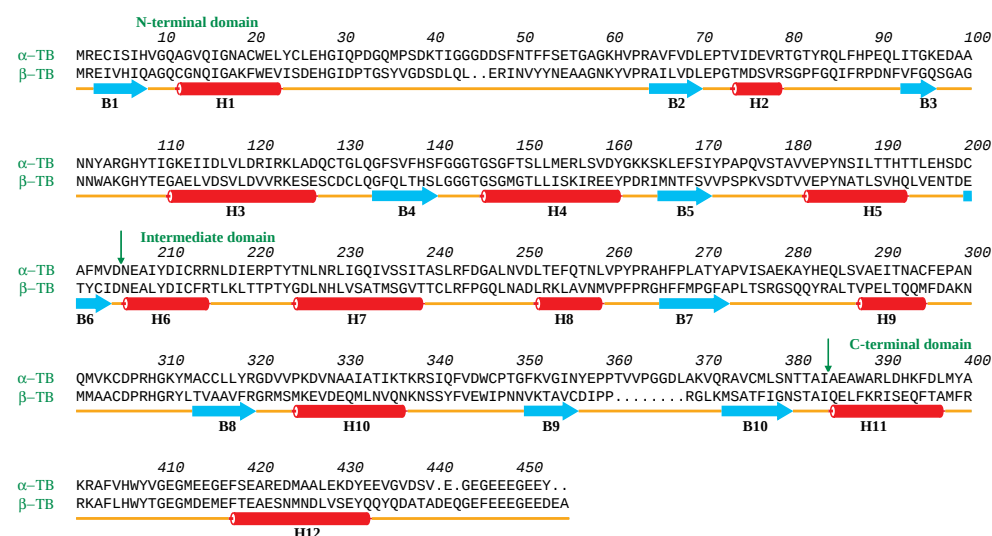


Figure 3 Sequences of pig brain α - and β -tubulin²⁸ used in the model (in the absence of tubulin sequences from cow we have used its closest known relative). Secondary structure elements are indicated and labelled as for Fig. 4. The tubulin preparations used in our experiments contained a mixture of isotypes. Most of the differences between isotypes are located at the extreme C terminus, which is not visible in our density. In most of the other positions of isotype differences, we arbitrarily chose the residue most similar to the other monomer.

B5–H5 loop is near the ribose, and N206 in H6, and Y224 and N228 in H7 interact with the nucleotide base. The position of the nucleotide agrees with models, based on comparison with other nucleotide-containing proteins, that position the phosphates near the glycine-rich loop. It is consistent with photocrosslinking experiments that locate C12 in β -tubulin near the guanine base¹⁰, the peptide β 157–176 near the ribose¹¹, and β 65–79 by the γ -phosphate¹². The position is also consistent with mutations in the region β 105–111 affecting the hydrolysis of GTP¹³. Finally, C12 and

either C203 or C213 can be crosslinked in β -tubulin, but only in the absence of the nucleotide¹⁴. In our model, the nucleotide is between those residues, but the distances between the cysteines are larger than 9 Å. However, the position of H6, containing C213, could easily be affected by the absence of nucleotide, bringing that cysteine closer to C12.

The nucleotide in one monomer interacts with the next monomer at the longitudinal interface. The interfaces between consecutive monomers in our density map are too similar to allow us to

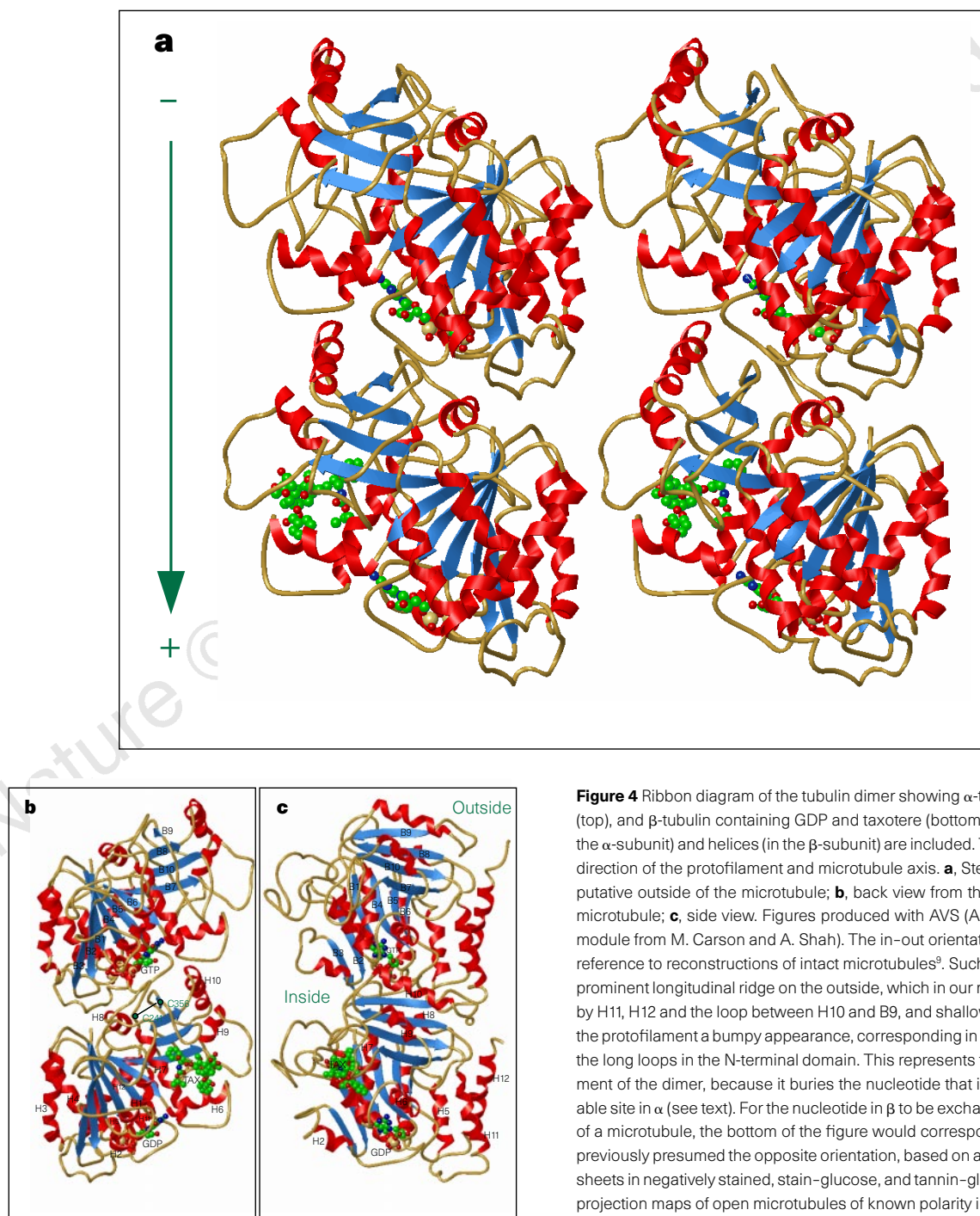


Figure 4 Ribbon diagram of the tubulin dimer showing α -tubulin with bound GTP (top), and β -tubulin containing GDP and taxol (bottom). Labels for strands (in the α -subunit) and helices (in the β -subunit) are included. The arrow indicates the direction of the protofilament and microtubule axis. **a**, Stereo front view from the putative outside of the microtubule; **b**, back view from the putative inside of the microtubule; **c**, side view. Figures produced with AVS (Advanced Visual; ribbon module from M. Carson and A. Shah). The in-out orientation was determined by reference to reconstructions of intact microtubules⁹. Such reconstructions show prominent longitudinal ridge on the outside, which in our model would be formed by H11, H12 and the loop between H10 and B9, and shallow inside grooves giving the protofilament a bumpy appearance, corresponding in our model to H1, B3 and the long loops in the N-terminal domain. This represents the most likely arrangement of the dimer, because it buries the nucleotide that is at the non-exchangeable site in α (see text). For the nucleotide in β to be exchangeable at the plus end of a microtubule, the bottom of the figure would correspond to the plus end. We previously presumed the opposite orientation, based on a comparison of the zinc sheets in negatively stained, stain-glucose, and tannin-glucose embedding, with projection maps of open microtubules of known polarity in negative stain⁹. Some ambiguity in that determination may be introduced by uncertainty about the exact rotational alignment of the protofilament in the sheets with respect to those in open microtubules and by stain artefacts. The polarity with the plus end down would be consistent with experiments that located the β -subunit at the plus end of the microtubule²⁹ and the α -subunit at the minus end³⁰. Circles in **b** indicate the positions of β Cys 241 and β Cys 356, separated by about 8 Å.

distinguish unequivocally between inter- and intradimer contacts (the interface near the N site is only slightly tighter than that by the E site). On the basis of the difference of exchangeability of nucleotide in the α and β monomers, we identified the intradimer interface as that in which the N site in the α -subunit is buried. The GTP at the E site in the β -subunit is partially exposed in the dimer but would be buried in the microtubule, where it becomes non-exchangeable. The assignment of the dimer shown in Fig. 4 thus provides the simplest explanation for nucleotide exchangeability in tubulin. Further support comes from data on the binding of colchicine. The main binding site of this drug is in the β -subunit, residues C356 and C241 (ref. 15) and region 1–36 (ref. 16) having been identified as part of the binding site (see the position of these residues in Fig. 4b). In addition, colchicine binds close to the $\alpha\beta$ intradimer interface¹⁷. On the other hand, antibodies against β 241–256, near the top of the β -monomer (Fig. 4b), and to α 214–226, at the bottom of the α -monomer, bind to the dimer but are unable to bind to the microtubule¹⁸. An alternative dimer to that shown in Fig. 4 would more readily explain these results, although the present model would be compatible if changes in the dimer occurred on binding to the microtubule.

On the basis of the similar behaviour of our sheets and microtubules, we believe that hydrolysis at the E site occurs upon sheet polymerization. In agreement with this, we find extra density in α -tubulin that corresponds to the γ -phosphate. The similarity between the two monomers, in spite of their being in two different nucleotide-bound states, is not surprising. The current model is that the GDP-containing dimer buried in the microtubule is kept locked in a GTP-like conformation by the microtubule lattice, and that only upon depolymerization does the GDP dimer 'spring' into a different conformation¹⁹. Comparison by electron microscopy of microtubules formed in the presence of GTP or of non-hydrolysable GTP analogues have shown only a change of 1.5 Å in the axial repeat (less than 2% of the subunit length)^{20,21}.

The favoured polymers for tubulin-GDP are rings formed by curved protofilaments²². The conformation of tubulin-GDP is thus referred to as 'curved', as opposed to the 'straight' conformation of tubulin-GTP (or tubulin-GDP held locked in the microtubule lattice)²³. In two simplified models, the curved conformation could either be one in which longitudinal contacts between dimers are at an angle (consistent with the dimer definition in Fig. 4), or one in which the dimer is bent at the monomer–monomer interface (alternative dimer definition). A more general model would involve extensive allosteric effects following hydrolysis, which could affect points distant from the nucleotide-binding pocket.

Our study was done on tubulin sheets that were stabilized by taxol. The model in Fig. 4 includes a molecule of taxol, a Taxol analogue whose structure has been solved by X-ray crystallography²⁴ (in taxol, the C-10 acetyl and the C-3' benzamide groups of Taxol have been replaced respectively by a hydroxyl and a N-t-BOC group). In our map, there is clear density for the taxane ring, the most dense and least flexible part of Taxol. Some density is also clear for one of the side chains. The position in the model is based on the best visual fit to the observed density. This position is in good agreement with photocrosslinking results placing the C-3' group near the sequence β : 1–31 (ref. 25), and the C2 group near the sequence β : 217–231 (ref. 26). In our model, the group at C-3' is near the top of helix H1 (that is, between β : 15–25), and the C2 group near H5 and the H5–H6 loop (that is, between β : 212–222). The main interaction of the taxane ring with tubulin is at L275, at the beginning of the B8–H9 loop.

Our model of tubulin shows a compact molecular structure with three functional domains: namely, GTP-binding, drug-binding and motor/MAP-binding domains. The interaction between domains is very tight, so the effects that nucleotides, drugs and other proteins in the cell have on tubulin are firmly linked. The assembly of tubulin

and its regulation through dynamic instability results from the fine tuning of the three components. Knowledge of the structure of tubulin should be invaluable for understanding the microtubule system in the cell.

The structure of the bacterial FtsZ protein is reported in this issue²⁷. Comparison of the tubulin and FtsZ models indicates that they have a common structural core of identical fold, which includes 10 β -strands surrounded by 10 α -helices: more detail will be revealed by careful comparison of these two structures. □

Methods

Crystalline tubulin sheets were polymerized in the presence of zinc from bovine brain tubulin (Cytoskeleton Inc.) and stabilized with taxol as described⁴. Samples were prepared by tannin–glucose embedding and examined at liquid-nitrogen temperature in a JEOL 4000 electron microscope at 400 kV following previously described procedures^{4,5}. Images were taken using spot-scan imaging and dynamic focus correction. Images and electron diffraction patterns were processed and merged as described^{4,9}.

Received 8 August; accepted 27 October 1997.

- Ludveña, R. F. The multiple forms of tubulin: different gene products and covalent modifications. *Int. Rev. Cyt.* **178**, 207–275 (1998).
- Mukherjee, A. & Lutkenhaus, J. Guanine nucleotide-dependent assembly of FtsZ into filaments. *J. Bacteriol.* **176**, 2754–2758 (1994).
- Gabor Miklos, G. L., Yamamoto, M., Burns, R. G. & Maleszka, R. An essential cell division gene of *Drosophila*, absent from *Saccharomyces*, encodes an unusual protein with tubulin-like and myosin-like peptide motifs. *Proc. Natl Acad. Sci. USA* **94**, 5189–5194 (1997).
- Nogales, E., Wolf, S. G., Zhang, S. X. & Downing, K. H. Preservation of 2-D crystals of tubulin for electron crystallography. *J. Struct. Biol.* **115**, 199–208 (1995).
- Nogales, E., Wolf, S. G., Khan, I. A., Ludueña, R. F. & Downing, K. H. Structure of tubulin at 6.5 Å and location of the taxol-binding site. *Nature* **375**, 424–427 (1995).
- Nogales, E., Wolf, S. G. & Downing, K. H. Visualizing the secondary structure of tubulin: three-dimensional map at 4 Å. *J. Struct. Biol.* **118**, 119–127 (1997).
- Burns, R. G. & Surridge, C. D. In *Microtubules* (eds Hyams, J. S. & Lloyd, C. W.) 3–32 (Wiley, New York, 1993).
- Little, M. & Ludueña, R. F. Structural differences between brain β 1- and β 2-tubulins: implications for microtubule assembly and colchicine binding. *EMBO J.* **4**, 51–56 (1985).
- Wolf, S. G., Nogales, E., Kikkawa, M., Grattinger, D., Hirokawa, N. & Downing, K. H. Interpreting a medium-resolution model of tubulin: comparison of zinc-sheet and microtubule structure. *J. Mol. Biol.* **263**, 485–501 (1996).
- Shivanna, B. D., Mejillano, M. R., Williams, T. D. & Himes, R. H. Exchangeable GTP binding site of β -tubulin—identification of cysteine 12 as the major site of cross-linking by direct photoaffinity labeling. *J. Biol. Chem.* **268**, 127–132 (1993).
- Hesse, J., Thierauf, M. & Ponstingl, H. Tubulin sequence region β 155–174 is involved in binding exchangeable guanosine triphosphate. *J. Biol. Chem.* **262**, 15472–15475 (1987).
- Linse, K. & Mandelkow, E.-M. The GTP-binding peptide of β -tubulin. Localization by direct photoaffinity labeling and comparison with nucleotide-binding proteins. *J. Biol. Chem.* **263**, 15205–15210 (1988).
- Davis, A., Sage, C. R., Dougherty, C. A. & Farrell, K. W. Microtubule dynamics modulated by guanosine triphosphate hydrolysis activity of β -tubulin. *Science* **264**, 839–842 (1994).
- Little, M. & Ludueña, R. F. Location of two cysteines in brain β 1-tubulin that can be cross-linked after removal of exchangeable GTP. *Biochim. Biophys. Acta* **912**, 28–33 (1987).
- Bai, R. *et al.* Identification of cysteine 354 of β -tubulin as part of the binding site for the A ring of colchicine. *J. Biol. Chem.* **271**, 12639–12645 (1996).
- Uppuluri, S., Knipling, L., Sackett, D. L. & Wolff, J. Localization of the colchicine-binding site of tubulin. *Proc. Natl Acad. Sci. USA* **90**, 11598–11602 (1993).
- Shearwin, K. E. & Timasheff, S. N. Effect of colchicine analogs on the dissociation of $\alpha\beta$ tubulin into subunits: the locus of colchicine binding. *Biochemistry* **33**, 894–901 (1994).
- Andreu, J. M. Site-directed antibodies to tubulin. *Cell Motil. Cytoskel.* **26**, 1–6 (1993).
- Caplow, M., Ruhlen, R. L. & Shanks, J. The free energy of hydrolysis of a microtubule-bound nucleoside triphosphate is near zero: all of the free energy for hydrolysis is stored in the microtubule lattice. *J. Cell Biol.* **127**, 779–788 (1994).
- Vale, R. D., Coppin, C. M., Malik, F. J. & Milligan, R. A. Tubulin GTP hydrolysis influences the structure, mechanical properties, and kinesin-driven transport of microtubules. *J. Biol. Chem.* **269**, 23769–23775 (1994).
- Hyman, A. A., Chrétien, D., Arnal, I. & Wade, R. H. Structural changes accompanying GTP hydrolysis of microtubules: information from a slowly hydrolyzable analog guanylyl-(α,β)-methylene-diphosphonate. *J. Cell Biol.* **128**, 117–125 (1995).
- Díaz, J. F., Pantos, E., Borda, J. & Andreu, J. M. Solution structure of GDP-tubulin double rings to 3 nm resolution and comparison with microtubules. *J. Mol. Biol.* **238**, 214–225 (1994).
- Mandelkow, E. & Mandelkow, E.-M. Microtubules and microtubule-associated proteins. *Curr. Opin. Cell Biol.* **7**, 72–81 (1995).
- Gueritte-Voegelein, F. *et al.* Structure of a synthetic taxol precursor: N-tert-butoxycarbonyl-10-deacetyl-N-debenzoyletaxol. *Acta Crystallogr. C* **46**, 781–784 (1990).
- Rao, S., Krauss, N. E., Heerding, J. M., Orr, G. A. & Horwitz, S. B. 3'-(p-Azidobenzamido)taxol photolabels the N-terminal 31 amino acids of β -tubulin. *J. Biol. Chem.* **269**, 3132–3134 (1994).
- Rao, S., Orr, G. A., Chaudhary, A. G., Kingston, D. G. I. & Horwitz, S. B. Characterization of the taxol binding site on the microtubule. *J. Biol. Chem.* **270**, 20235–20238 (1995).
- Löwe, J. Y. & Amos, L. A. Crystal structure of the bacterial cell-division protein FtsZ complexed with GDP. *Nature* **391**, 203–206 (1998).
- Ponstingl, H., Krauss, E., Little, M., Kempf, T., Hofer-Warbinek, R. & Ade, W. Amino acid sequence of α - and β -tubulins from pig brain: heterogeneity and regional similarity to muscle proteins. *Cold Spring Harbor Symp. Quant. Biol.* **46**, 191–197 (1982).
- Mitchison, T. J. Localization of an exchangeable GTP binding site at the plus end of microtubules. *Science* **261**, 1044–1047 (1993).

30. Fan, J., Griffiths, A. D., Lockhart, A., Cross, R. A. & Amos, L. A. Microtubule minus ends can be labeled with a phage display antibody specific to α -tubulin. *J. Mol. Biol.* **259**, 325–330 (1996).

Acknowledgements. We thank R. F. Ludueña for isotopically purified α II and α III tubulin, M. Le for help with electron diffraction processing, and R. M. Glaeser and Y. L. Han for comments on the manuscript. Taxol was provided by the Drug Synthesis and Chemistry Branch, Division of Cancer Treatment of the National Cancer Institute. This work was supported by the NIH.

Correspondence and requests for materials should be addressed to E.N. Coordinates referred to in this Letter have been deposited in the Brookhaven Protein Data Bank with ID 1tub and will be accessible within one year.

Crystal structure of the bacterial cell-division protein FtsZ

Jan Löwe & Linda A. Amos

MRC Laboratory of Molecular Biology, Cambridge CB2 2QH, UK

Bacterial cell division ends with septation, the constriction of the cell wall and cell membranes that leads to the formation of two daughter cells^{1,2}. During septation, FtsZ, a protein of relative molecular mass 40,000 which is ubiquitous in eubacteria and is also found in archaea and chloroplasts³, localizes early at the division site to form a ring-shaped septum. This septum is required for the mechanochemical process of membrane constriction⁴. FtsZ is a GTPase^{5,6} with weak sequence homology to tubulins⁷. The nature of FtsZ polymers *in vivo* is unknown, but FtsZ can form tubules, sheets and minirings *in vitro*^{8,9}. Here we report the crystal structure at 2.8 Å resolution of recombinant FtsZ from the hyperthermophilic methanogen *Methanococcus jannaschii*. FtsZ has two domains, one of which is a GTPase domain with a fold related to one found in the proteins p21^{ras} and elongation factor EF-Tu. The carboxy-terminal domain, whose function is unknown, is a four-stranded β -sheet tilted by 90° against the β -sheet of the GTPase domain. The two domains are arranged around a central helix. GDP binding is different from that typically found in GTPases and involves four phosphate-binding loops and a sugar-binding loop in the first domain, with guanine being recognized by residues in the central connecting helix. The three-dimensional structure of FtsZ is similar to the structure of α - and β -tubulin¹⁰.

Two FtsZ genes (named after filamenting temperature-sensitive mutant Z) from the archaeon *M. jannaschii* have been characterized by the genome project¹¹. One gene, MJ0370, was amplified by genomic polymerase chain reaction (PCR) and expressed in *E. coli*/C41, a mutant of BL21 capable of expressing toxic genes¹². Proteolysis during cell disruption was minimized by using heat-shock treatment. Cubic crystals were obtained and the structure was solved by multiple isomorphous replacement and density modification (see Methods and Table 1). The model (Fig. 1) contains residues 23–356, 116 water molecules, and one molecule of GDP; weak density for residues 1–22 was visible as an extension from helix H0.

FtsZ consists of two domains with a long, 23-residue, helix H5 (Figs 1a, 2) connecting them. The N-terminal portion of the molecule, containing residues 38–227, has GDP obtained from the expression host bound to it and will be called the GTPase domain. It consists of a six-stranded parallel β -sheet surrounded by two and three helices on both sides. The overall fold of the GTPase domain of FtsZ is related to typical GTPases and can be superimposed on the p21^{ras}–GDP complex (Protein Data Bank (PDB) entry 1Q21; ref. 13) using 52 C α atoms (S1, H1, S2, H2, S4, H3 and S5) to give a root-mean-squared (r.m.s.) deviation of 1.88 Å. The topology of the β -sheet in FtsZ is 321456, which is slightly different from the topology in p21^{ras} (ref. 13), where it is 231456, but,

together with the arrangement of five helices (H1, HL1, H2, H3 and H4), is consistent with typical Rossmann-fold topology¹⁴. Helix H2A is unique to FtsZ. Numbering of secondary structure elements (Fig. 2) follows the corresponding elements of p21^{ras} proteins.

The C-terminal domain, spanning residues 228–356, consists of a mainly parallel four-stranded central β -sheet supported by two helices on one side. The topology of the sheet is 1423, with strand 4 antiparallel to the others. The uncovered side of the sheet makes contacts with helix H5 and is otherwise open to the solvent. The fold of the C-terminal domain is related to chorismate mutase of *Bacillus subtilis* and can be superimposed on PDB entry 1COM¹⁵ with an r.m.s. deviation of 1.83 Å over 52 C α atoms (SC1, HC2, SC2, HC3, SC3 and SC4). Additionally, sequence comparisons give similarities to calmodulins in three loop regions (Swissprot CALM-TRYCR; loops between H5/HC1, SC1/HC2, and SC2/HC3) and to adenyl cyclase (CYA1_HUMAN; residues 620–740), making a role in calcium binding feasible. The electrostatic potential on the

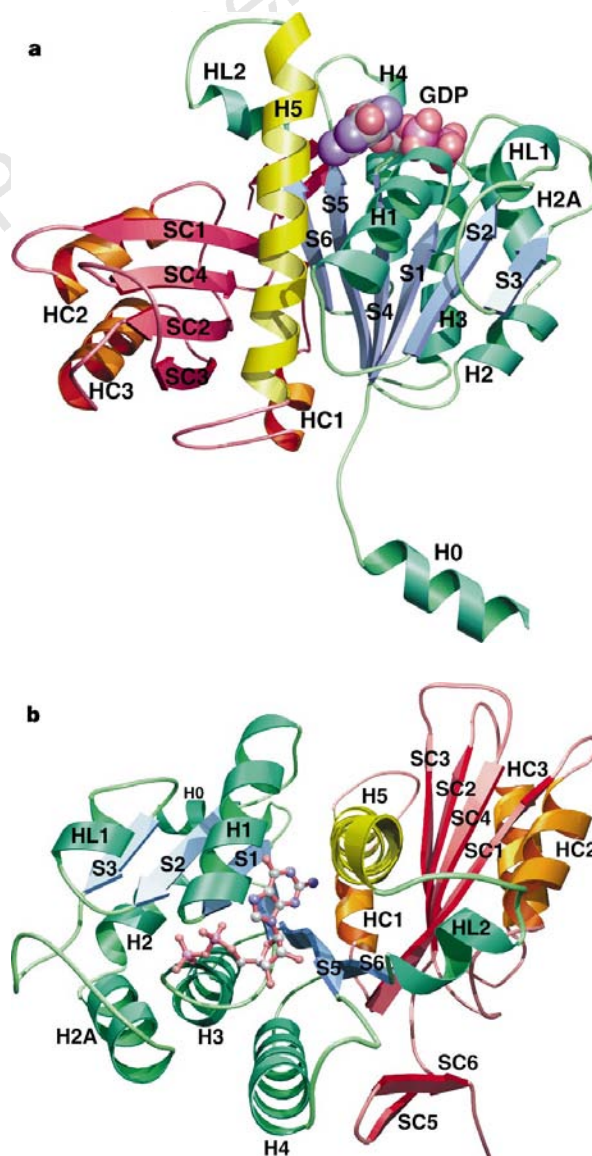


Figure 1 Ribbon drawings of FtsZ (residues 23–356) from *M. jannaschii*. **a**, View showing the GTPase domain in blue/green, the C-terminal domain in red/orange, and the connecting helix H5 in yellow. GDP is represented by a space-filling model. **b**, View of FtsZ rotated by ~90° from that in **a**. GDP is represented by a ball-and-stick model. Figures were prepared with POVSCRIPT (D. Peisach, personal communication)²⁸.

surface of the C-terminal domain is dominated by a large patch of acidic residues on the open side of the β -sheet. This region forms crystal contacts to partly disordered residues 1–10 of a symmetry-related molecule. Residues 343–352 at the C terminus form a small β -hairpin which contacts S5 and H4 of the GTPase domain. Residues 357–372 are disordered in the crystal. The electron density for the C-terminal domain is slightly weaker than for the GTPase domain, probably because the C-terminal domain has only a few crystal contacts.

As predicted^{16,17}, GDP binding to FtsZ is different from typical GTPases¹⁸, although the GTPase domain of FtsZ is related to the fold of typical GTPases like p21^{ras}. Six distinct sequence regions in FtsZ make contacts to the nucleotide: loops T1–T4 (tubulin loops) and two regions for sugar binding and guanine binding, respectively (Fig. 2). The GTPase domain starts with the typical strand–loop–helix motif between S1 and H1, but the conformation of this loop is different from that in classical P-loop proteins. FtsZ makes three additional contacts to the nucleotide's phosphates: loop T2 between

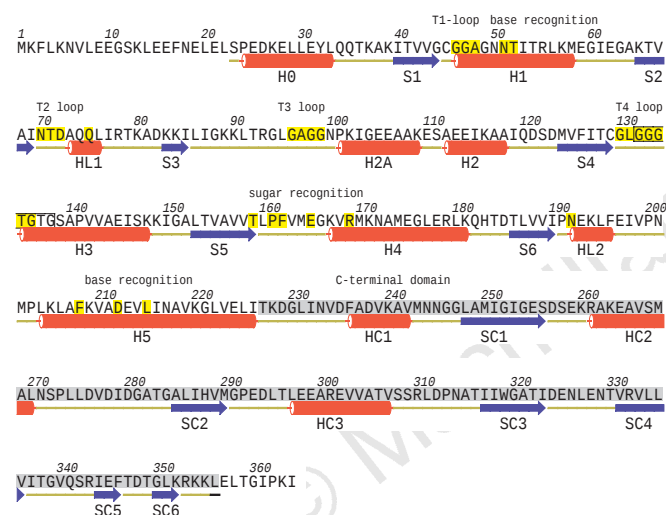


Figure 2 Secondary structure assignment of FtsZ1 from *M. jannaschii*. The secondary structure of FtsZ determined from the crystal structure was calculated with DSSP²⁹. The numbering of secondary structure elements follows the p21^{ras} nomenclature¹³. Elements in the C-terminal domain are numbered SC and HC for strands and helices, respectively. Residues highlighted in yellow make contact with GDP. Residues of the C-terminal domain are shaded. Prepared using ALSRIPT³⁰.

S2 and HL1, loop T3 between S3 and H2A, and loop T4 between S4 and H3 carrying the tubulin signature motif. All four loops make mainly backbone contacts to the phosphates. As FtsZ was crystallized in the presence of EDTA, we did not expect to find a magnesium ion complexed to GDP. A poorly ordered water molecule (WAT472) occupies a similar position to the magnesium ion in p21^{ras}–GDP¹³; however, WAT482 is complexed between the β -phosphate and Asn 70, which makes it unlikely that this position in FtsZ would be occupied by a positively charged magnesium ion. Furthermore, it has been reported that nucleotide binding to FtsZ, but not nucleotide hydrolysis, is independent of magnesium¹⁹. The structure displays a pocket for the γ -phosphate and we found that fresh crystals showed weak electron density for a γ -phosphate. We think it is possible that GTP, originally bound to the protein, is slowly hydrolysed in the crystals. We were not able to detect large conformational changes between new and old crystals but the molecule may be fixed in one state by crystal restraints. The loop between S5 and H4 contains two residues that specifically bind the sugar moiety of the nucleotide (Fig. 3): Glu 165 hydrogen-bonds with the two hydroxyl groups O2' and O3' and Arg 169 makes contacts to the O5' hydroxyl and the α -phosphate. Guanine recognition is mainly accomplished by residues within H5: Asp 212 points to N1 and Phe 208 stacks on the aromatic ring. This aromatic stack is extended by Phe 162 and Phe 196. Leu 215 makes a hydrophobic contact on the other side of the guanine.

Table 1 Crystallographic data

	NAT12	NAT11	Thiomersal (3 mM)	
Space group / 2,3 (199)			EMTS1 (48 h)	EMTS3 (18 h)
Temperature	100 K	RT	RT	RT
Resolution (Å)	2.8	3.7	4.0	4.2
Completeness (%)	99.0 (99.7)	97.3	90.2	97.1
R_{merge}^*	0.065 (0.31)	0.12	0.14	0.15
$I/\sigma(I)$ last shell	2.3	1.7	1.6	2.1
$R_{\text{iso}}^\dagger$			0.19	0.16
Phasing power ‡			1.88	1.83
Refinement				
Model	Residues 23–356, 1 GDP and 116 water molecules			
Data	NAT12, 8–2.8 Å, all reflections			
R factor	0.199 (R_{free} 0.282§)			
R.m.s. deviations	Bonds: 0.012 Å, angles: 1.77°; temperature factors: 3.4 Å ²			

Completeness and R_{merge} for the outermost shell (2.80–2.95 Å) are in parentheses for NAT12. * $R_{\text{merge}} = \sum_i \sum_j |I(h,i) - \langle I(h) \rangle| / \sum_i \sum_j I(h,i)$ where $I(h,i)$ are symmetry-related intensities and $\langle I(h) \rangle$ is the mean intensity of the reflection with unique index h . $^\dagger R_{\text{iso}} = 2\sigma_h / (F_{\text{PH}} - F_P) / (\sigma_h / F_P + F_{\text{PH}})$ where F_{PH} and F_P are the derivative and native structure factor amplitudes, respectively. ‡ Phasing power: mean value of the heavy-atom structure amplitudes divided by the residual lack-of-closure. Figure of merit was 0.59 and 0.42 for centric and acentric reflections, respectively. § Five per cent of reflections were selected for determination of free R factor.

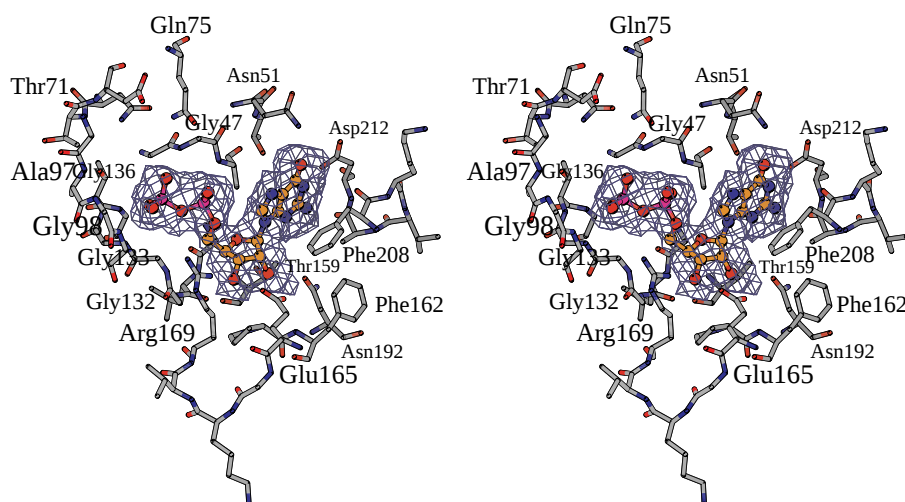


Figure 3 Stereo plot of the active site of FtsZ. Superimposed is the final $2F_o - F_c$ electron density map for the nucleotide contoured at 1σ . Prepared with MOLSCRIPT²⁸.

Asn 51 forms a hydrogen bond with O6 of the base. By contacting H5, the nucleotide may induce a conformational change during hydrolysis, moving the relative orientation of the catalytic domain and the C-terminal domain. A region around Gly 221 may deliver the flexibility between the two β -sheets for this type of movement. As in other GTPases, no residue could be identified that provides an activated nucleophile for hydrolysis. A comparison between the active sites of FtsZ and p21^{ras} in the GDP-bound state¹³ reveals important differences: the P-loop in p21^{ras} (G1 box, residues 10–17) has no counterpart in FtsZ but is probably functionally related to loops T1 and T4 in FtsZ. Residues in T1 and T4 make backbone contacts to the phosphates (residues 44–48 and 130–136, containing the tubulin signature motif). Thr 135 of FtsZ is approximately in the same position as Ser 17 of p21^{ras}. The T2 loop in FtsZ (residues 70–75) is approximately in the same place as the residues of the G2 box of p21^{ras} (residues 32–37). Asp 72 of FtsZ is the only acidic side chain in the active site of FtsZ and may be the equivalent of Asp 57 of p21^{ras}. Loop T3 in FtsZ (residues 94–99) is nearly in the same place as G box G3 of p21^{ras} (residues 57–60) with respect to the phosphates. Residues Lys 117 and Asp 119, located in a loop between S5 and S6, make the most important contacts to the guanine ring in p21^{ras}. The equivalent residues in FtsZ are Phe 208 and Asp 212, respectively. These residues are located in helix H5 in FtsZ. The sugar moiety in FtsZ is rotated by $\sim 180^\circ$ with respect to p21^{ras} if the central β -sheet of the two proteins is aligned. The residue making the hydrogen bonds to O2' and O3' in p21^{ras} is located in the first half of the sequence (p21^{ras} Asp 30), whereas the corresponding residue in FtsZ (Glu 165) is located in the second half of the GTPase domain in the sugar-binding loop (residues 159–169). Sequence comparisons between tubulins and FtsZ had already revealed limited homology of ~ 10 –18% identity^{7,19,20}. Sequence alignment based on the structure of FtsZ alone (data not shown) revealed that most residues involved in GDP binding and several structural glycines and prolines are conserved and made it likely that the GTPase domain of tubulin is a Rossmann fold as well. The C-terminal domain of tubulins is larger and has very limited sequence homology to FtsZ, possibly reflecting diverse differences between the variety of interacting molecules. Strong structural similarity between FtsZ and tubulin both in the GTPase and the C-terminal domain is now evident from comparing the structure of FtsZ with the two-dimensional crystal structure of α - and β -tubulin¹⁰. Tubulin lacks the N-terminal extension of FtsZ but contains a large insertion between H1 and S2. Two large helices in tubulin extend the C-terminal domain. Nucleotide binding is very similar between FtsZ and β -tubulin, with many residues involved in GDP binding being conserved.

The 22 partly disordered N-terminal residues stick out of the molecule and make important crystal contacts, as do the last visible residues at the C terminus, which form a twofold contact with symmetry-related residues of another molecule. Inspection of the crystal packing reveals no obvious protofilament structures, as expected for a crystal packing with one molecule per asymmetric unit and a high-symmetry space group. The exposed N terminus with ~ 33 residues sticking out of the compact GTPase domain is rapidly degraded by proteases in the C41 expression host. Comparison between different FtsZ sequences reveals that this N-terminal extension is highly variable and is very short in *E. coli*, where it starts at residue 28, so it is unlikely to be important in filament formation. The C terminus is visible up to residue 356, with eight residues of the wild-type protein and the additional eight residues of the His₆-tagged protein being disordered. The C termini of FtsZ sequences are divergent from residue 342 onwards, after the last sheet of the C-terminal domain. Both the C and N termini contain many acidic residues and acidic/hydrophobic repeats. These sequences could simply enhance the solubility of monomers or filaments, although other functions have been proposed for the C termini of tubulins²¹. FtsZ(MJ0370) from *M. jannaschii* contains two cysteines, Cys 129

and Cys 45, which lie near the active-site strands S1 and S4 in the GTPase domain: this geometry would allow formation of a disulphide bridge and one mercury atom can bind to both residues. Because of its high similarity to tubulin, FtsZ should prove to be a simple model system for microtubule dynamics and be amenable to protein engineering. □

Methods

Protein expression, purification and crystallization. Genomic DNA was extracted from a living culture of *M. jannaschii* (DSM 2661, ATCC 43067) with a commercial kit (Promega Wizard). Genomic PCR using two primers (5'-GAAGTCCATATGAAATTCTAAAAACGTTTAA-3', 5'-CGTATTAGGATCCAATTTTGGAAATTCCTGTGTGTCTA-3'), replacing the GTG start codon with ATG, produced a 1,115-bp DNA carrying the complete 1,092 bp of MJ0370 and unique cleaving sites for *Nde*I and *Bam*HI. This fragment was cloned into the pHis17 vector (B. Miroux, personal communication), putting it under control of a T7-promoter and adding eight residues at the C terminus (GSHHHHHH) to yield a protein of 372 residues. C41 cells¹² were transformed and expressed the His-tagged protein after addition of IPTG. After induction in log phase, cells from a 10-litre culture were collected and snap-frozen in liquid nitrogen. The frozen pellet was powdered under liquid nitrogen and poured directly into 200 ml of boiling buffer A (50 mM Tris-HCl, 300 mM NaCl, pH 8.0) and stirred for 90 s. After addition of ice to a final volume of 400 ml, the lysate was centrifuged and applied to a Ni-NTA column (Qiagen). The FtsZ protein was eluted at ~ 400 –450 mM imidazole in buffer A, pH 6.0. The protein was further purified on a Sepharose-6B column (Pharmacia) using 20 mM Tris-HCl, 1 mM EDTA, 1 mM NaN₃. FtsZ eluted as an oligomer under low-salt conditions. The product was checked by electrospray mass spectrometry (observed: 39,889.12 g mol⁻¹; estimated: 39,891.33 g mol⁻¹). Crystals were grown by the hanging-drop vapour-diffusion technique using 0.1 M MES pH 6.62, 17% PEG400 and 6% ethanol as crystallization solution. Droplets composed of 3 parts of 10 mg ml⁻¹ protein solution and 1 part crystallization solution were equilibrated for a minimum of one week. Two different crystal forms grew under these conditions, distinguishable by the lack of birefringence for one of them. These latter crystals belong to cubic spacegroup *I*2₁3, with one molecule per asymmetric unit and 71% solvent content; cell dimensions are $a = b = c = 159.14$ Å, with all angles 90° . Crystals for multiple isomorphous replacement were collected in crystallization solution and mounted in sealed capillaries. A crystal for the frozen dataset NATI2 was collected in crystallization solution and soaked for 5 min in 0.1 M MES pH 6.62, 25% PEG200 and 6% ethanol and frozen in a stream of cold nitrogen at 100 K.

Structure determination and refinement. Each low-resolution dataset was collected from a single crystal, using a MAR imaging-plate detector mounted on an Elliot GX13 rotating anode X-ray generator. The final native dataset NATI2 was collected from a frozen crystal at beamline 9.6 at the Daresbury SRS. Images were indexed and integrated with the MOSFLM²² package and further processed using the CCP4 suite of programs²³. The structure was determined by the MIRAS method using only one derivative (ethylmercurithiosalicylic acid, thiomersal, 3 mM) with two different soak times of 18 and 48 h, respectively. Heavy-atom positions were determined using SHELXS²⁴ and initial phases were calculated with MLPHARE. Phasing statistics are given in Table 1. After 20 cycles of phase refinement using SOLOMON²⁵ and phase extension from 4.0 to 3.7 Å resolution, a superb electron density map was obtained. Residues 1–356 could be built into the first density using FRODO²⁶, although the density for residues 1–22 was weak and these residues were later omitted from the final model. Crystallographic refinement was performed using the programs REFMAC (room-temperature data sets; phase-restrained) and X-PLOR²⁷. A total of 26 cycles of refinement and manual rebuilding was necessary to switch from the unfrozen to the frozen native dataset. At 2.8 Å resolution, solvent mask correction was applied and a resolution range of 8.0–2.8 Å was used. No reflections were excluded from refinement by their signal-to-noise ratio. The current model contains residues 23–356, 116 water molecules and 1 GDP molecule. Residues 1–22 gave weak difference density but were not refineable. The model shows good geometry with an r.m.s. deviation over bond length of 0.012, bond angles of 1.77° and no Ramachandran outliers. Individual temperature factors have been refined using tight restraints. They are higher for the C-terminal domain (residues 227–356), with an overall temperature

factor over residues 23–356 of 53 Å² (65 Å² from the Wilson plot). Coordinates and structure factors have been deposited in the Brookhaven Protein Data Bank.

Received 4 August; accepted 20 October 1997.

- Rothfield, L. I. & Justice, S. S. Bacterial cell division: the cycle of the ring. *Cell* **88**, 581–584 (1997).
- Donachie, W. D. The cell cycle of *Escherichia coli*. *Annu. Rev. Microbiol.* **47**, 199–230 (1993).
- Erickson, H. P. FtsZ, a tubulin homologue, in prokaryote cell division. *Trends Cell Biol.* **7**, 362–367 (1997).
- Lutkenhaus, J. FtsZ ring in bacterial cytokinesis. *Mol. Microbiol.* **9**, 403–409 (1993).
- de Boer, P., Crossley, R. & Rothfield, L. The essential bacterial cell-division protein FtsZ is a GTPase. *Nature* **359**, 254–256 (1992).
- Raychaudhuri, D. & Park, J. T. *Escherichia coli* cell-division gene FtsZ encodes a novel GTP-binding protein. *Nature* **359**, 251–254 (1992).
- Erickson, H. P. FtsZ, a prokaryotic homolog of tubulin? *Cell* **80**, 367–370 (1995).
- Bramhill, D. & Thompson, C. M. GTP-dependent polymerization of *Escherichia coli* FtsZ protein to form tubules. *Proc. Natl Acad. Sci. USA* **91**, 5813–5817 (1994).
- Mukherjee, A. & Lutkenhaus, J. Guanine nucleotide-dependent assembly of FtsZ into filaments. *J. Bacteriol.* **176**, 2754–2758 (1994).
- Nogales, E., Wolf, S. G. & Downing, K. H. Structure of the $\alpha\beta$ tubulin dimer by electron crystallography. *Nature* **391**, 199–203 (1998).
- Bult, C. J. *et al.* Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **269**, 496–512 (1996).
- Miroux, B. & Walker, J. E. Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **260**, 289–298 (1996).
- Tong, L., de Vos, A. M., Milburn, M. V. & Kim, S.-H. Crystal structures at 2.2 Å resolution of the catalytic domains of normal ras protein and an oncogenic mutant complexed with GDP. *J. Mol. Biol.* **217**, 503–513 (1991).
- Rossmann, M. G., Moras, D. & Olsen, K. Chemical and biological evolution of a nucleotide-binding protein. *Nature* **250**, 194–199 (1974).
- Chook, Y. M., Gray, J. V., Ke, H. & Lipscomb, W. N. The monofunctional chorismate mutase from *Bacillus subtilis*: structure determination of chorismate mutase and its complexes with a transition state analog and prephenate, and implications on the mechanism of enzymatic reaction. *J. Mol. Biol.* **240**, 476–500 (1994).
- Linse, K. & Mandelkow, E.-M. The GTP-binding peptide of β -tubulin. *J. Biol. Chem.* **263**, 15205–15210 (1988).
- Sage, C. R. *et al.* Site-directed mutagenesis of putative GTP-binding sites of yeast β -tubulin: evidence that α -, β - and γ -tubulins are atypical GTPases. *Biochemistry* **34**, 16870–16875 (1995).
- Bourne, H. R., Sanders, D. A. & McCormick, E. The GTPase superfamily: conserved structure and molecular mechanism. *Nature* **349**, 117–127 (1991).
- Mukherjee, A., Dai, K. & Lutkenhaus, J. *Escherichia coli* cell division protein FtsZ is a guanine nucleotide binding protein. *Proc. Natl Acad. Sci. USA* **90**, 1053–1057 (1993).
- de Pereda, J. M., Leynadier, D., Evangelio, J. A., Chacon, P. & Andreu, J. M. Tubulin secondary structure analysis, limited proteolysis sites, and homology to FtsZ. *Biochemistry* **35**, 14203–14215 (1996).
- Sackett, D. L. Structure and function in the tubulin dimer and the role of the acidic carboxyl terminus. *Subcell. Biochem.* **24**, 255–302 (1995).
- Leslie, A. G. W. Recent changes to the MOSFLM package for processing film and image plate data (SERC Laboratory, Daresbury, Warrington, UK, 1991).
- The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D* **54**, 760–763 (1994).
- Sheldrick, G. M. Heavy atom location using SHELXS-90. In *Isomorphous Replacement and Anomalous Scattering: Proceedings of the CCP4 Study Weekend 25–26 January 1991* (eds Wolf, W., Evans, P. R. & Leslie, A. G. W.) 23–38 (SERC Daresbury Laboratory, Warrington, UK, 1991).
- Abrahams, J. P. & Leslie, A. G. W. Methods used in the structure determination of bovine mitochondrial F1-ATPase. *Acta Crystallogr. D* **52**, 30–42 (1996).
- Jones, T. A. A graphics model building and refinement system for macromolecules. *J. Appl. Cryst.* **11**, 268–272 (1978).
- Brünger, A. X-PLOR Version 3.1, A System for X-ray Crystallography and NMR (Yale University Press, New Haven and London, 1992).
- Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946–950 (1991).
- Kabsch, W. & Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983).
- Barton, G. J. ALSICRIPT: a tool to format multiple sequence alignments. *Prot. Eng.* **6**, 37–40 (1993).

Acknowledgements. We thank H. Huber (University of Regensburg) for *M. jannaschii* cells, I. Fearnley for mass spectrometry, B. Miroux for C41 cells and for discussion, and A. Murzin for pointing out the similarity to chorismate mutase. J.L. is supported by an EMBO long-term fellowship.

Correspondence and requests for materials should be addressed to J.L. The identity codes for the FtsZ coordinates and structure factors in the Protein Data Bank are 1FSZ and 1RFSZF, respectively.

YOURS TO HAVE AND TO HOLD BUT NOT TO COPY

The publication you are reading is protected by copyright law. Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA. Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931 Fax: 0171 436 3986

Twisting your robotic arm

This roundup pulls together products from the areas of automation and DNA amplification, including barcoding LIMS systems, robotic sample processors, microplate carousel systems, PCR kits and thermal cyclers.

High-throughput organic synthesis

From Gilson

A new eight-page brochure featuring this personal organic synthesizer system has been recently been released

The brochure explains how this integrated, automated chemical synthesis system utilizes the high-throughput organic synthesis technique to accelerate drug discovery. Designed for both lead generation and lead optimization libraries, the system can handle both solution- and solid-phase combinatorial chemistry techniques. It is well suited to this area because its reaction vessel racks can be easily removed from the workstation during incubation, while the system maintain both temperature and atmospheric control. In addition, the system can be configured with up to eight reaction blocks from a variety of manufacturers. According to the manufacturer, time-consuming manual transfers are eliminated — cleaved material can be transferred directly into a microplate rather than into collection tubes, eliminating an additional liquid transfer step and hence material loss before purification, analysis or characterization. The brochure also highlights the advanced informatics/control software from Afferent Systems.

Reader Enquiry No. 100

Genesis robotic sample processor

From Tecan

A system designed for automated sample preparation in molecular biology laboratories

This robotic sample processor (RSP) is well suited for automated processing of a range of molecular biology applications, including DNA extraction, preparation of DNA sequencing reactions and polymerase chain reaction (PCR). It is a liquid handling system that pipettes into tubes or microplates with 6- to 384-well formats. It can be used in conjunction with a robotic manipulator arm, transporting microplates or other con-



Tecan's Genesis robotic sample processor.

tainers around the worktable, as well as to and from stackers. Custom-built systems can also incorporate equipment like thermal cyclers or nucleic acid extraction systems positioned on the instrument. Disposable tips with integral filters help to minimize the risk of carryover contamination and help to ensure reliability in sensitive applications, such as PCR.

Reader Enquiry No. 101

MDS 2.5

From Sagian

This is a fully featured, application-control software package for the ORCA laboratory robotics system

ORCA automates repetitive lab protocols to increase sample throughput and productivity. MDS 2.5 expands the capabilities of the control program and adds Windows95 compatibility. Improvements to this release include enhanced speed and graphics, expanded access to and manipulation of application dictionaries, as well as the ability to search for and replace text across multiple procedure entries. MDS 2.5 supports up to nine COM peripheral ports. It also allows quick programming of the ORCA arm, including multi-tasking. Moreover, the program data can be integrated with a wide variety of third-party software programs. The MDS 2.5 software upgrade will be available free-of-charge to all support contract customers. It may be purchased directly through Sagian in the United States and Canada and through Beckman Instruments in other countries.

Reader Enquiry No. 102

SkannerStacker

From Skatron

Intended for washing and stacking 96-well microplates, this unit should help to provide improved high-throughput screening

An internal pressure pump is used to set individual pressure settings and accommodates applications for cell washing and the Coombs test, as well as standard ELISAs. The system features two inlets, flexible programming for optimal assay results, and 'prime' and 'rinse' programmes, making the washer well suited for most applications and needs. The SkannerStacker includes two magazines, each holding 25 microplates, and a transfer unit for easy release and loading of microplates. The unit can be programmed to restack the microplates before or after washing, thus creating a 'first in, first out' situation. According to Skatron, the pro-



Rack'em and stack'em — SkannerStacker performs washing and stacking of 96-well microplates.

cessing time is 15 seconds per plate, including wash time, which helps to make the SkannerStacker an efficient washer.

Reader Enquiry No. 103

WinLIMS

From QSI

A new barcoding facility creates and reads barcodes for this laboratory information management system

Barcoded information can now be read directly into LIMS and automatically placed in its correct tabular format within the system. Barcoded information reduces the potential for keystroke errors and drastically improves input speed. Sample data can be input at any time via a hand-held scanner. Descriptions are included with information of when, where and who received the sample, as well as previous test references and even hazard information. Other members of staff can be left messages of where the sample is physically located and different shifts and levels of personnel can quickly pick up the sample checking and testing procedure without disruption. Barcodes can be produced from any point within WinLIMS, but usually from the designated 'sampling point', where labelling routines can be activated for the new batches taken. Barcoded information is automatically placed into the relevant field within WinLIMS and updated information need only be entered in one field for all the rest to be instantly brought into accord. The barcoding facility is available as standard on all new WinLIMS packages or as an upgrade for existing users.

Reader Enquiry No. 104

Bar-One Platinum

From Zebra Technologies

A barcode labelling system with integrated software

Version 4.0 of this software utilises the Windows operating system, allowing users to design and print custom labels for a wide range of applications. Users can access, process and combine complex, variable information, including barcodes, text and graphics from multiple data sources and initiate print jobs from almost any legacy or client/server application. The new software also has improved networking capabilities and comes with a five-seat user license. Bar-One Platinum 4.0 meets the open database connectivity (ODBC) standard from Microsoft, enabling users to share data with any ODBC-compliant application. Bar-One features a graphical drag-and-drop ODBC facility to help users establish relationships between databases, sorts and filters to create variable and customized label formats quickly and easily. New object alignment and spacing tools improve the speed of design, and a preview feature for the print queue allows users to see the label before loading it for printing. Bar-One is designed to optimize the power and performance of Zebra printers through ZPL II.

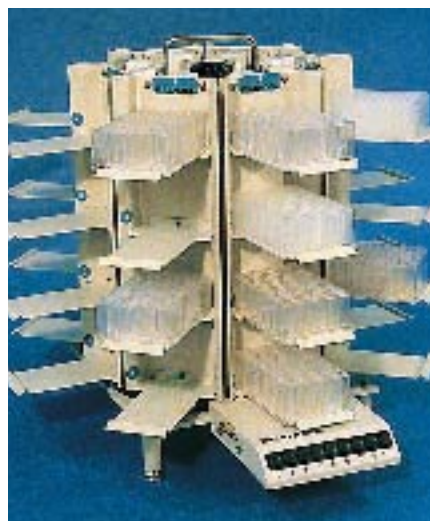
Reader Enquiry No. 105

The Carousel

From Brandel

For use with robotic arms, this high-capacity plate carousel is capable of holding up to 160 standard 96-well plates

These storage systems are compatible with most robotic arms and are well suited for use with the company's robotic harvester or the Beckman Biomek 2000 side-loader system. The Carousel consists of eight individual shelf stacks, which can be removed and altered — each is capable of holding up to 20 standard-size, 96-well plates. Interchangeable columns allow for combinations of plates, blocks and mini-tubes. Carousels can be supplied for storing plates either lengthwise or widthwise and customized configurations are also available. Two-way communication, using an RS-232 interface, enables the carousel to be programmed to turn either way, with controlled, indexed rotation, bringing each of the stacks to a common location for use with the robotic arm. It will also, if required, agitate the plates at varying shaking speeds and degrees of oscillation, heat and cool the plates, and even introduce CO₂ (it can be operated manually). A custom system for use with the Beckman's Biomek 2000 side-loader system features a carousel that can hold either 64 plates, 32 tip boxes, 32 reservoirs or 24 P1000 tip boxes. It has a robotic system that can hold 192 microtitre plates,



Riding the Carousel, a new system from Brandel.

more than tripling Biomek's capacity when three carousels are installed together.

Reader Enquiry No. 106

BioGrid

From BioRobotics

A smaller size robot gives smaller overall size higher throughput to this micro-array deposition technology, which is due to be released in the first quarter of 1998

Gridding and library replication are now made faster and simpler with this robot, which is able to grid 36 plates onto 24 filters in 90 min, with only one operator intervention. It uses 60 per cent less bench space than other currently available gridding instruments. Set-up time is also decreased by using a built-in BioBank loading system, which changes and loads source plates in a few seconds. The risk of contamination is reduced as the lid is only removed from individual source plates during sampling. Designed with no moving electrical components, the instrument's construction improves reliability and reduces laboratory noise. In addition, the use of a free-floating, square-headed design prevents rotation of the picking pins. As the BioGrid is designed to work alongside a colony picking instrument, such as the BioPick, both operations can be carried out simultaneously.

Reader Enquiry No. 107

APS 300



Bio Grid has micro-array deposition technology.

From Don Whitley Scientific

A compact automated power stacker that offers rapid hands-off plate preparation

Developed by AES Laboratoire the system features a single carousel that has a capacity for 300 90-mm Petri dishes. A rapid pouring rate of 600 agar plates per hour and automatic stacking provides efficient 'hands-free' plate preparation for busy laboratories. Operation and programming is guided by straightforward questions and prompts displayed on the LCD. The plate carousel has convenient loading and its mechanically driven transfer disc can be easily removed for cleaning. Once a programme has been initiated, the economical APS 300 becomes a 'walk-away' system, requiring no further intervention, says the manufacturer. Inverted plates are simply ignored and stacked empty without any interruption to operation. The APS 300 can store up to 40 programmes and an optional printer allows the generation of reports detailing the agar name, batch number, date, operator identification, volume prepared and number of dishes filled. Other optional extras include a facility to print onto the dishes, bi-plate filling and refrigeration. For maximum convenience, the 'easy-load' peristaltic pump is automatically calibrated and allows the system to be used manually for filling tubes and bottles. A different version of the unit, the APS 300/55 provides automatic stacking of 55–60-mm plates.

Reader Enquiry No. 108

Plato 7

From Rosys Anthos

Plato 7 combines up to eight-tip liquid handling with robotic plate manipulation

Optional modules for plate incubation, shaking, washing and reading can be added to the system, as required, allowing more automation of the user's protocol. Plate stacking modules also offer high plate-loading capacities. With a new gantry arm for greater positional accuracy and new pipetting technology, this system lends itself to higher density plate formats. Windows95 or NT control software provides an open, comfortable, intuitive user interface to programme customized applications. Also available is Lucy 1: a high-quality, photo-optics system that offers a microplate reader with some additional talents. Lucy 1 is said to combine lumino-metric and photometric analysis on a variety of microplate formats. With precision internal reagent injectors and Peltier temperature control of the measuring chamber, single measurements or complex kinetic sequences may be carried out. Lucy 1 has its own graphic display and keyboard or it can be controlled from a PC with



Contemplate the Plato 7 liquid handling system.

Windows software for flexible and convenient set-up.

Reader Enquiry No. 109

Amplifier blowout

SunRise

From Oncor

A system for simultaneous amplification and detection of nucleic acids within the PCR tube

The method is based on the incorporation of energy-transfer-labelled hairpin 'primers' into the amplified DNA product. Using Sunrise UniPrimer I, the system can perform amplification of virtually any target. It consists of a hairpin primer designed such that a fluorescent signal is generated only when the primer is unfolded by incorporation into the amplification product. Unincorporated UniPrimer I has virtually no fluorescent signal. Most importantly, this eliminates the need to purify the PCR product before detection. The closed-tube format of the Sunrise system eliminates PCR contamination, post-PCR sample manipulation (such as gel analysis) and easily accommodates high-throughput analysis of amplified material.

Reader Enquiry No. 110

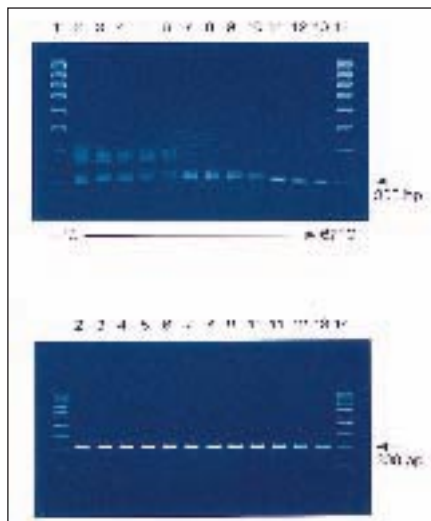
Blunt-end PCR cloning kit

From Boehringer Mannheim

A single kit that can be used to clone both small or large DNA fragments

This kit applies Boehringer's DNA ligation system to achieve fast, efficient cloning of small PCR DNA fragments (<1.5 kb), as well as larger fragments of up to 10 kb. The 'trick', claims the manufacturer, lies in the suicide vector (pCAPs) containing the lethal mutant gene of the catabolite activator protein (CAP). Ligation of the blunt-ended DNA fragment, generated by PCR or other methods, disrupts the expression of the CAP gene, so only positive recombinants grow after transformation. Blunt-ended DNA fragments can be used directly for blunt-end ligation to achieve high cloning efficiencies without the need for pretreatment, such as purification or polishing. With a stated cloning efficiency of >95 per cent, the kit should prove useful to those using the Expand High Fidelity System, or thermostable DNA polymerases, such as Pwo DNA polymerase, to generate blunt ends.

Reader Enquiry No. 111



Don't feel blue about contamination — Eppendorf offers shielded PCR test tubes.

Shielded PCR test tubes

From Eppendorf

A shield on the lid of these 0.2-ml PCR test tubes prevents accidental contact with the inner surface of the lid during handling

The contamination shield is designed to make opening and closing of this miniature test tube less of a chore. The test tubes have a special hinge, which allows the lid to click into a set position of approximately 90°, preventing it from covering adjacent tubes in the thermal block. These PCR tubes are available individually, as well as in five- and eight-well strips. Accessories include work trays in the microtitre plate format, which can be used to position the test tubes in the 96-well thermal block. The work tray is combined with a frame to form a test tube rack, so that these miniature test tubes are easy to handle.

Reader Enquiry No. 112

Gene Pool cDNA

From Invitrogen

PCR-ready, high-quality, first-strand cDNA

Gene Pool cDNA is available from a variety of human tumour and normal tissues, eliminating the need to acquire tissue or isolate RNA. Gene Pool cDNA can be used to amplify a gene of interest with specific or degenerative primers, analyse tissue or tumour-specific gene expression, or characterize alternately spliced mRNA.

Reader Enquiry No. 113

Quanti-Path kits

From CPG

Ultra-sensitive non-isotopic detection of pathogen-specific PCR products in 96-well format

These kits are designed for ease of use and require less than 60 min of hands-on time. The results are reproducible and give greater

sensitivity, with more than 10 copies of target sequence. The detection procedure is performed at room temperature with no need to run gels. The kits include a 96-well plate coated with capture oligonucleotides that bind to PCR amplicons generated using the primers included with each kit. Detection is accomplished by the simple colorimetric assay provided. Quantification is attained by optimization of PCR cycling conditions and titration of the target standard. These kits are available for detection of human immunodeficiency virus (HIV-1 and HIV-1,2), hepatitis B virus, hepatitis C virus, human cytomegalovirus, human herpesvirus (HHV-6), herpes simplex viruses (HSV-1 and HSV-1,2), Varicella-Zoster virus, and *Borrelia burgdorferi*.

Reader Enquiry No. 114

RT-PCR primer pairs

From Promega Neurosciences

Neurotrophic factors and receptors for use in RT-PCR procedures

The manufacturer is now offering eight sets of optimized primer pairs for analysing expression of specific mRNAs: BDNF, CNTF, GDNF, NGE, TrkB, p75/LNGFR, ChAT and beta-actin by RT-PCR. Most of the primer pairs span introns and, therefore, can be used to distinguish amplification of cDNA versus genomic DNA. The

ADVERTISEMENT

"Beside good science, there is much humour and much of interest outside the field of science. ... In my case, it kept me awake and delighted until the small hours"

Keith J. Laidler, Professor Emeritus of Chemistry, Ottawa University

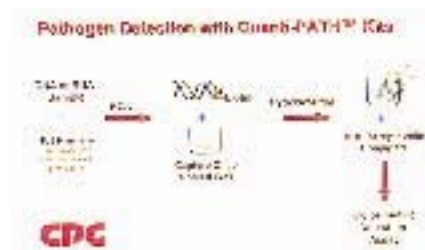
The Bedside Book Genius and Eccentricity in Science 1869–1953

Edited by Walter Gratzer, with a foreword by Stephen Jay Gould

A fascinating selection of articles from Nature, offering a panorama of science seen against the background of 19th and 20th century history.



Price: £19.95/US\$29.95 + postage
Discount rate available for current subscribers.
To order, or for more information, contact: Nature Marketing
Tel. +44 (0) 171 843 4985
Fax. +44 (0) 171 843 4998
E-mail: subscriptions@nature.com



CPG's non-isotopic pathogen detection in a 96-well format.

► primers have been tested with total human and rat brain RNA in RT-PCR. They are provided in separate tubes containing 100 µM each, which is sufficient for twenty RT-PCR reactions. These primer pairs are the first in a series of oligonucleotide primers for research into neurotrophic factors.

Reader Enquiry No. 115

Mini-Gene

From Stuart Scientific

An economical thermal cycler that has been designed for improved accuracy and versatility with smaller samples

The reaction block is made from precision-drilled aluminium for both robustness and uniform heat conductivity. Blocks are available for 14 × 0.5-ml or 31 × 0.2-ml microcentrifuge tubes, as well as an 8 × 4 format microtitre plate. Available as an option is the Iso-Temp heated lid, which provides for oil-free cycling. Mini-Gene uses a temperature probe located within the reaction block to relay the actual sample temperature to the microprocessor. Continual monitoring of the sample temperature ensures that the programmed temperature profile is accurately and reproducibly achieved within all sample vessels. The system's software is flexible and user-friendly, prompting the user at each stage of programming. The auto-control system helps to ensure that there are controlled ramping times. The wide-angle, back-illuminated screen gives constant updates on elapsed time, temperature and cycle number during operation. Finally, an RS232 interface allows connection to a computer or printer.

Reader Enquiry No. 116



Iso-Temp lid for the Mini-Gene thermal cycler.

XX-15S bench lamp

From UVP

A shortwave (254 nm), ultraviolet radiation lamp for sterilization that reduces bacterial counts in laboratory areas

As sterilization in biological laboratories is of great importance, the XX-15S is a good tool for giving shortwave ultraviolet coverage of 1.4 mm² at 50 cm from the lamp fixture. The lamp can be used in an upward facing position for purifying the air, or in a downward position for sanitizing work surfaces. The metal lamp housing is designed with internal aluminum reflectors for maximum ultraviolet exposure. Brackets are included for ceiling or wall mounting. The lamp may also be hung freeing benchtop space.

Reader Enquiry No. 117

GeneAmp optical reaction plate

From PE Applied Biosystems

A 96-well plate engineered specifically for the GeneAmp PCR System 9600 and the ABI Prism 7700 sequence detection system

The plate incorporates a built-in tray to ensure that the heated lid will have the correct pressure on the MicroAmp caps or MicroAmp whole-plate covers. The built-in tray also provides a barrier to ambient air and helps to ensure temperature accuracy and uniformity. The new plate is made of a proprietary polypropylene material used in MicroAmp tubes and caps. The dimensions of the tube section are the same as regular MicroAmp tubes, and the shape and thickness of the tubes is accurate to ±0.025 mm. The outer tube walls are frosted to improve performance in sequence detection instruments. Tube sealing is accomplished using either MicroAmp strip caps or whole-plate covers, and the plates are packaged in individually sealed bags in boxes of ten plates.

Reader Enquiry No. 118

PCR product guide

From Robbins Scientific

An free, updated guide to the company's line of PCR plasticware

Among the products described in this 12-page, colour brochure are Split Well CyclePlate-384 and CyclePlate-192, the next generation of thin-walled PCR plates designed for high-throughput PCR and sequencing applications. In addition to the original 96-sample CyclePlate, the company now offers CyclePlate-24 and CyclePlate-8 for handling fewer samples per run. Updated information on StripEase PCR tubes and RoboTips for use on Beckman Instruments' Biomek laboratory workstation is also included.

Reader Enquiry No. 119

RoboCycler Infinity

From Stratgene

A temperature cycler with 40-, 96- and 384-well interchangeable blocks

The newly released RoboCycler Infinity cycler combines features of the standard RoboCycler temperature cycler with the added flexibility of interchangeable blocks. Most conventional temperature cyclers have only one block for PCR denaturation, annealing and extension steps and must ramp a single block's temperature up and down. Every RoboCycler temperature cycler features four programmable thermal blocks and a robotic arm that rapidly moves samples from one pre-equilibrated block to another. This is said to result in a 30 per cent reduction in cycling time. Another key specification of the system is ±0.1 °C, well-to-well temperature uniformity. The new RoboCycler Infinity temperature cycler adds the versatility of easily interchangeable blocks. An independent Hot Top assembly for the system allows oil-free PCR amplifications.

Reader Enquiry No. 120

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

Peptides
\$18/residue

Antibodies
monoclonal - \$8500
polyclonal - \$1195

Complete services for your custom peptide and polyclonal or monoclonal antibody needs.

Research Genetics, Inc.
U.S. or Canada 800-533-4363
U.K. 0-800-89-1393
FAX 205-536-9016

READER ENQUIRY NO. 121

Bacterial & Yeast Growth Media

Over 400 recipes and formulae in stock

BIO 101

BIO 101

BIO 101 Media is available in capsule, powder & pouch form

Contact BIO 101 or your local distributor for ordering & information
(800) 424-6101 • fax: (760) 598-0116
web: www.bio101.com • e-mail: info@bio101.com

READER ENQUIRY NO. 122